

PSIHOLOŠKA OBZORJA

*Slovenska znanstveno-strokovna psihološka revija
(uradna revija Društva psihologov Slovenije)*

Glavna in odgovorna urednica

doc. dr. Anja Podlesek

HORIZONS OF PSYCHOLOGY

Official Journal of the Slovenian Psychological Association

Editor-in-Chief

Assist. Prof. Dr. Anja Podlesek

Izdaja / Published by

Društvo psihologov Slovenije

(Slovenian Psychological Association)

Ljubljana

ISSN 1318-1874

Letnik (Volume) 17, številka (Number) 3, 2008

PSIHOLOŠKA OBZORJA HORIZONS OF PSYCHOLOGY

Slovenska znanstveno-strokovna psihološka revija
Official Journal of the Slovenian Psychological Association
Letnik (Volume) 17, številka (Number) 3, 2008, ISSN 1318-1874

Izdaja / Published by
Društvo psihologov Slovenije / Slovenian Psychological Association
Ulica Stare pravde 2, 1000 Ljubljana, Slovenija

Glavna in odgovorna urednica / Editor-in-Chief
ANJA PODLESEK
Univerza v Ljubljani, Filozofska fakulteta
Oddelek za psihologijo
Aškerčeva 2, 1000 Ljubljana, Slovenija
Tel. +386 (0)1 241 1184, Fax: +386 (0)1 42 59 301
email: anja.podlesek@ff.uni-lj.si

Uredniški odbor / Editorial Board

VALENTIN BUCIK
LUKA KOMIDAR
DARE KOVAČIČ
TATJANA SAMEC
MOJCA VIZJAK PAVŠIČ
URŠKA ŽUGELJ

Uredniški svet / Scientific Board

TIZIANO AGOSTINI
KARIN BAKRAČEVIČ VUKMAN
GAŠPER CANKAR
ROBERT CVETEK
DARJA KOBAL GRUM
BRIGITA KRIČAJ KORELC
LIDLIA MAGAJNA

VLADO MIHELJAK
ALJOSCHA NEUBAUER
SONJA PEČJAK
CVETA RAZDEVŠEK PUČKO
ARGIO SABADIN
META SHAWE-TAYLOR

VERA SLODNJAK
GREGOR SOČAN
IRENA ŠINIGOJ BATISTIČ
LEA ŠUGMAN BOHINC
VLADIMIR TAKSIČ
PREDRAG ZAREVSKI
MAJA ZUPANČIČ

Lektoriranje angleških ključnih besed, vsebinska obdelava in razvrstitev besedila: *Urška Žugelj*

Oblikovanje ovitka: *Jana Leskovec*

Tehnično urejanje in prelom: *Luka Komidar*

Tisk: *Tiskarna Vovk d.o.o., Domžale*

Pogostost izhajanja: *Na leto izidejo štiri številke*

Ustanovitelj, izdajatelj in založnik: *Društvo psihologov Slovenije*

Naročniški naslov: *Psihološka obzorja, Društvo psihologov Slovenije, Ulica Stare pravde 2, 1000 Ljubljana, Slovenija, telefon/fax: +386 (0)1 282 1086*

Davčna številka: SI 51264218; Poslovni račun: 02010-0091342732

© 2008 Društvo psihologov Slovenije. Revija je zaščiten z zakonom o avtorskih pravicah. Revija v celoti ali katerikoli njen del ne smeta biti razmnoževana ali posredovana na noben način (fotokopije, mikrofilmi, elektronsko posredovanje, snemanje, prevajanje) brez predhodnega pisnega privoljenja izdajatelja.

Bibliografski zapisi o prispevkih, objavljenih v reviji *Psihološka obzorja*, so ustrezno kategorizirani in vključeni v slovensko podatkovno zbirko COBIB. Revija je opredeljena kot pomembna in odmevna znanstvena in strokovna periodična publikacija v psihološki znanosti, torej so objave v njej visoko referenčne za napredovanje v znanstvenih in strokovnih nazivih. **Psihološka obzorja so indeksirana v PsycINFO, svetovni bibliografski bazi psihološke literature.**

Revija izhaja s finančno podporo Agencije za raziskovalno dejavnost Republike Slovenije ter Znanstvenoraziskovalnega inštituta Filozofske fakultete Univerze v Ljubljani.

Vsebina

Znanstveni raziskovalnoempirični in teoretsko-pregledni prispevki

<i>Špela PROGAR in Gregor SOČAN</i> <i>An empirical comparison of Item Response Theory and Classical Test Theory</i>	5
<i>Andrej KASTRIN</i> <i>Metaanaliza in njen pomen za psihološko metodologijo</i>	25
<i>Maja ZUPANČIČ, Helena R. SLOBODSKAYA in Gennady G. KNYAZEV</i> <i>Gender differences in the child/adolescent personality traits: Slovenes and Russians compared</i>	43
<i>Peter PRAPER</i> <i>From chaos through cosmos toward coinonia: A group-analytic developmental line</i>	65
<i>Igor AREH</i> <i>Povezava med Eysenckovimi osebnostnimi potezami in spominsko obnovo</i>	75
<i>Branko SLIVAR</i> <i>Ugotavljanje vzorca stresorjev pri delu učiteljev v povezavi z zadovoljstvom pri delu</i>	93

Recenzije knjig

<i>Melita PUKLEK</i> <i>Recenzija znanstvene monografije "Psihološki vidiki preventivnih dejavnosti v prometu"</i>	113
<i>Polona MATJAN ŠTUHEC</i> <i>Ocena knjige "Skupinska psihoterapija: Od mitologije do teorije"</i>	117
<i>Predvideni prispevki v številki 4 letnika 17 (2008) PSIHOLOŠKIH OBZORIJ</i>	120
<i>Navodilo avtorjem prispevkov</i>	123

Contents

Scientific papers (empirical research and theoretical papers/reviews)

<i>Špela PROGAR and Gregor SOČAN</i> <i>An empirical comparison of Item Response Theory and Classical Test Theory</i>	5
<i>Andrej KASTRIN</i> <i>Meta-analysis: Its role in psychological methodology</i>	25
<i>Maja ZUPANČIČ, Helena R. SLOBODSKAYA and Gennady G. KNYAZEV</i> <i>Gender differences in child/adolescent personality traits: Slovenes and Russians compared</i>	43
<i>Peter PRAPER</i> <i>From chaos through cosmos toward coinonia: A group-analytic developmental line</i>	65
<i>Igor AREH</i> <i>Connection between Eysenck's personality traits and memory recall</i>	75
<i>Branko SLIVAR</i> <i>Examination of the pattern of teachers' work stressors in relation to job satisfaction</i>	93

Book reviews

<i>Melita PUKLEK</i> <i>Review of the book "Psychological aspects of prevention in traffic"</i>	113
<i>Polona MATJAN ŠTUHEC</i> <i>Review of the book "Group psychotherapy: From mythology to theory"</i>	117
<i>Content of the next issue, No. 4, Vol. 17 (2008), of HORIZONS OF PSYCHOLOGY</i>	120
<i>Instructions for contributors</i>	125

An empirical comparison of Item Response Theory and Classical Test Theory

Špela Progar¹ and Gregor Sočan^{2}*

¹*Mirna Peč, Slovenija*

²*University of Ljubljana, Department of Psychology, Ljubljana, Slovenia*

Abstract: Based on nonlinear models between the measured latent variable and the item response, item response theory (IRT) enables independent estimation of item and person parameters and local estimation of measurement error. These properties of IRT are also the main theoretical advantages of IRT over classical test theory (CTT). Empirical evidence, however, often failed to discover consistent differences between IRT and CTT parameters and between invariance measures of CTT and IRT parameter estimates. In this empirical study a real data set from the Third International Mathematics and Science Study (TIMSS 1995) was used to address the following questions: (1) How comparable are CTT and IRT based item and person parameters? (2) How invariant are CTT and IRT based item parameters across different participant groups? (3) How invariant are CTT and IRT based item and person parameters across different item sets? The findings indicate that the CTT and the IRT item/person parameters are very comparable, that the CTT and the IRT item parameters show similar invariance property when estimated across different groups of participants, that the IRT person parameters are more invariant across different item sets, and that the CTT item parameters are at least as much invariant in different item sets as the IRT item parameters. The results furthermore demonstrate that, with regards to the invariance property, IRT item/person parameters are in general empirically superior to CTT parameters, but only if the appropriate IRT model is used for modelling the data.

Key words: item response theory, classical test theory, measurement invariance, psychometrics

Empirična primerjava teorije odgovora na postavko in klasične testne teorije

Špela Progar¹ in Gregor Sočan²

¹*Mirna Peč*

²*Univerza v Ljubljani, Oddelek za psihologijo, Ljubljana*

Povzetek: Teorija odgovora na postavko (TOP) temelji na modelu odnosa med latentno lastnostjo in odgovorom na postavko ter posledično omogoča neodvisno ocenjevanje parametrov postavk in oseb ter lokalno oceno napake merjenja. Kljub tem teoretičnim prednostim TOP v primerjavi s klasično testno teorijo (KTT) pa rezultati empiričnih raziskav pogosto ne pokažejo sistematičnih razlik med parametri po KTT in TOP ter med stopnjo invariantnosti parametrov po KTT in TOP. V pričujoči empirični študiji smo skušali, na podlagi podatkov iz Mednarodne raziskave trendov znanja iz matematike in

*Naslov/Address: doc. dr. Gregor Sočan, Univerza v Ljubljani, Oddelek za psihologijo, Aškerčeva 2, 1000 Ljubljana, tel.: +386 1 241 11 84, fax.: +386 1 42 59 301, e-mail: gregor.socan@ff.uni-lj.si

naravoslovja (TIMSS 1995), odgovoriti na naslednja raziskovalna vprašanja: (1) Kako primerljivi so parametri postavk in oseb po KTT in TOP? (2) Kako invariantni so parametri postavk KTT in TOP preko različnih skupin oseb? (3) Kako invariantni so parametri postavk in oseb po KTT in TOP preko različnih skupin postavk? Ugotovili smo, da so parametri oseb in postavk po KTT in TOP zelo primerljivi, da so parametri postavk po KTT in TOP podobno invariantni v različnih skupinah oseb, da so parametri oseb po TOP bolj invariantni preko različnih skupin postavk ter da so parametri postavk po KTT, ocenjeni iz različnih skupin postavk, v enaki meri ali celo bolj invariantni od parametrov postavk po TOP. Rezultati so pokazali tudi, da so na splošno parametri postavk po TOP, v smislu invariantnosti parametrov, empirično superiorni parametrom po KTT, vendar le v primeru, ko se uporabljeni model TOP v zadostni meri sklada s podatki.

Ključne besede: teorija odgovora na postavko, klasična testna teorija, merska invariantnost, psihometrija

CC = 2200

Item-response theory (IRT) appears to be the currently prevailing paradigm within the psychometric theory. However, this is only partially reflected in the psychometric practice: with an important exception of educational measurement, most psychological measuring instruments still appear to be based on the classical test theory (CTT). The aim of this study was to examine some possible causes of this gap, besides conservatism of the psychometric practitioners. We tried to provide a partial answer to two commonly asked questions: “To what extent does the choice of the test construction method influence the properties of the resulting test?” and “To what extent are the theoretical advantages of the IRT reflected in empirical superiority of the resulting measurement instruments?” In the sequel, we shall first review some important properties of both paradigms and summarize the available empirical evidence.

The main advantage of CTT is its simplicity: the estimation of the measured trait is performed simply by adding the scored responses to items. CTT does not involve truly latent variables: despite the fact that the true score is not empirically observable, it can be defined operationally as the average score on the infinite number of equivalent repetitions of the measurement process (see, for instance, Lord and Novick, 1968). Accordingly, within the CTT framework, the question of model validity is almost never addressed. The basic equation of CTT, the additive decomposition of the observed score into the true score and the random error, is a tautology rather than a model and can not possibly be subjected to a test. Groups of assumptions, called “measurement models” (most important among them being the congeneric and the tau-equivalent model, respectively) do play a role in CTT, but they often affect optimality rather than validity: for instance, if the test items depart largely from the essential tau-equivalence, coefficient alpha will remain the lower bound to reliability, but will be a relatively inefficient lower bound. On the other hand, the assumptions which do affect the validity of the results—for instance, the

assumption of uncorrelated errors—are mostly impossible to be tested empirically for practical reasons. Finally, some parts of CTT are based on a model which is incorrect *a priori*: for instance, although the reliability can be defined as the squared correlation coefficient between the true and observed scores, this relationship is not linear almost surely whenever the range of possible scores is bounded (which is almost always the case).

On the other hand, IRT is a model-based paradigm: it starts with modelling the relationship between the latent variable being measured and the item response. The aim is, at least in principle, to find an accurate model rather than a robust approximation. The question of model fit therefore always plays an important role in an IRT analysis. An important feature of the modelling approach is that the parameters of the persons do not depend on parameters of the items, and vice versa. Consequently, parameters of the persons are invariant across items, and parameters of the items are invariant in different populations of persons. However, these properties depend on the validity of the item-response model. Since real data never perfectly fit to some reasonably parsimonious abstract model, questions can arise about how large the lack of fit can be so that one can still rely on these properties, and whether IRT is always superior to CTT with regard to accuracy and invariance of parameter estimates. Little evidence is available so far regarding these questions, partly perhaps because the model fit has traditionally been treated dichotomously rather than in sense of the degree of fit.

Apart from the invariance problem, another relevant question is whether one should expect IRT and CTT to produce substantially different tests when selecting items from a larger pool. A common feature of both approaches is to discard items to which the answers are not related, or are even negatively related to the test score. In classical test construction (see, for instance, Nunnally and Bernstein, 1994), the primary item selection criterion is the corrected item-total correlation. Additionally, the spread of item difficulties should be large enough so that the discrimination power is not concentrated around the average scores only. Because of the stress on the item discrimination indices, classical tests can be expected to have both relatively high internal consistency and high unidimensionality¹; the former may be quantified by means of coefficient alpha or some other similar reliability coefficient, and the latter by means of the proportion of common variance, explained by the first common factor. On the other hand, IRT test analysis does not involve the notion of reliability, at least not as defined in CTT. So an IRT-based test may not have a very high internal consistency; instead, it will not tend to concentrate the discrimination power around average score – in fact, the concentration of the discrimination power can be set according to the aim of the test constructor. In any case, somewhat different products should be expected depending on the paradigm used in the process of test construction.

¹ It should be noted that the relationship between reliability and unidimensionality is intricate; for a discussion, see Ten Berge and Sočan (2004).

Recently, several studies compared CTT and IRT, mostly with regard to the comparability of the item and person parameters. For instance, Fan (1998) in his study, based on real data, found very high correlations both among the person parameters (correlations were in all instances higher than .96) and among the item difficulties (correlations higher than .90). He further found no evidence of a higher invariance of the IRT item parameters in comparison to the CTT item parameters. Similar conclusions were reached by Courville (2005) in another empirical study.

An obvious limitation of empirical studies is that the parameter values can not be manipulated and that the true values are not known. MacDonald and Paunonen (2002) conducted a Monte Carlo study in which they controlled the spread of item difficulty and item discrimination. Similarly to Fan (1998) and Courville (2005), they found high correlations between the CTT and the IRT difficulties. The discrimination indices, however, correlated highly only when the spread of discriminations was large and the spread of difficulty values was small. Moreover, the CTT discrimination estimates were in some conditions (i.e., at a large spread of difficulties) less accurate than the IRT estimates.

The abovementioned studies did not attempt to construct a test by selecting items from a larger item pool and to subsequently investigate the invariance of the parameters across such item sets, although this may be a fairly typical situation in practice. Our study was therefore set into the item selection framework. Within this framework, it attempted to address the following standard questions:

- How comparable are CTT and IRT based item and person parameters?
- How invariant are CTT and IRT based item parameters across different participant groups?
- How invariant are CTT and IRT based item and person parameters across different item sets?

Method

Data source

The data used in our study are from the Third International Mathematics and Science Study (TIMSS 1995) administered in 1994 and 1995 to the third and fourth grade students in 45 different countries. Mathematical achievement was measured with 102 and science achievement with 97 different items, respectively. Items were divided in 26 exclusive clusters, which were distributed into 8 different test booklets. Most of the items (approx. 80%) were multiple-choice items, while others required short or more elaborated answers. Complete international database is available on <http://isc.bc.edu/timss1995i/Database.html>.

Since test booklets consisted of different clusters of items, each student responded just to a subset of items, which represents a problem for data analysis using CTT. Therefore only items from two different test booklets were selected and analyzed for the purpose of this study, i.e. items from booklet number 5 for the mathematical achievement and booklet number 6 for the science achievement. Booklet number 5 consisted of 39 mathematical items (two items were eliminated, because they were not administered in Latvia) and booklet 6 consisted of 37 science items as shown in Table 1. These items represent math and science item pools. Items were dichotomous, except for three mathematical items and one science item, which were subsequently dichotomized for the purpose of this study.

Table 1. *Structure of mathematical and science item pools.*

	Math Item Pool	Science Item Pool
Multiple-choice items	25	26
Short answer items	7	3
Elaborated answer items	7	8
Total	39	37

Participants

From the complete international database only students from six European countries (Hungary, Latvia, Netherlands, Norway, Scotland, and Slovenia) that responded to items from booklet 5 and 6 were selected. The structure of students that responded to mathematical and science items is shown in Tables 2 and 3.

Table 2. *Structure of students that responded to mathematical items (from booklet number 5).*

	Lower grade				Upper grade			
	Female		Male		Female		Male	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Hungary	175	18.7	199	19.5	189	19.6	182	18.3
Latvia	125	13.3	126	12.3	135	14.0	143	14.4
Netherlands	152	16.2	201	19.7	162	16.8	152	15.3
Norway	134	14.3	132	12.9	132	13.7	149	15.0
Scotland	199	21.2	192	18.8	197	20.5	213	21.5
Slovenia	152	16.2	171	16.7	148	15.4	153	15.4
Total	937	100.0	1021	100.0	963	100.0	992	100.0

Note. 35 students for which gender was not indicated are not included in this table.

Table 3. Structure of students that responded to science items (from booklet number 6).

	Lower grade				Upper grade			
	Female		Male		Female		Male	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Hungary	195	20.2	175	18.0	187	19.4	182	18.2
Latvia	130	13.5	127	13.1	139	14.4	137	13.7
Netherlands	172	17.9	176	18.1	162	16.8	156	15.6
Norway	134	13.9	142	14.6	124	12.9	150	15.0
Scotland	178	18.5	195	20.0	199	20.6	214	21.4
Slovenia	154	16.0	158	16.2	153	15.9	163	16.3
Total	963	100.0	973	100.0	964	100.0	1002	100.0

Note. 49 students for which gender was not indicated are not included in this table.

Data analysis

Data from both the math and the science item pool were analyzed using both CTT and IRT procedures. Missing value analysis was conducted using EM (expectation-maximization) algorithm in *SPSS 11.0 for Windows* before item analysis (the percentage of students that had at least one missing response was 0.3 % in the math item pool and 0.5 % in the science item pool).

Within the CTT framework the following parameters were computed:

- person parameter as the proportion of correct answers,
- item discrimination parameter as the corrected point-biserial correlation,
- item difficulty parameters as the proportion of correct responses to particular items and
- alpha reliability coefficient and the greatest lower bound to reliability (GLBR). The latter was computed according to the algorithm proposed by Ten Berge, Snijders and Zegers (1981).

The greatest lower bound to reliability (GLBR) is the highest value which is certainly not higher than the actual reliability in the sample. It is therefore the most accurate conservative reliability estimate possible. GLBR does not have a closed-form solution and can only be estimated by numerical algorithms.

The IRT analysis was conducted by means of *BILOG-MG 3.0* (Zimowski, Muraki, Mislevy, & Bock, 1996). All IRT estimations were obtained using the marginal maximum likelihood (MML) method with normal prior distribution, which is the default for *BILOG-MG*.

Within the IRT framework the following parameters were computed:

- person parameter (commonly known as the *theta* value),
- item discrimination (slope) parameter (the *a* value),

- item difficulty (location) parameter (the b value).

Information about the item and model fit was obtained through the differences in log likelihoods between different unidimensional logistic models (1PL, 2PL, and 3PL models) and through comparisons between expected and empirical item characteristic curves (ICC's; χ^2 test and the graphical method). Unidimensionality of the tests was analyzed with minimum rank factor analysis (MRFA; Ten Berge and Kiers, 1991), which is the single available factor analytic method making possible the determination of the proportion of the common variance, explained by some number of common factors. In contrast to this, other similar methods only make possible to evaluate the proportion of the total variance, explained by the common factors, which is however not directly associated with unidimensionality. In our case, the proportion of the common variance of the items, explained by the first common factor, was taken as a measure of unidimensionality.

Comparability of CTT and IRT person and item parameters

Comparability of the CTT and the IRT person parameters was assessed by correlating the proportion correct and theta values for participants in both math and science item pool. Comparability of the CTT and the IRT item parameters was assessed by correlating the CTT and the IRT based item difficulty parameters and item discrimination parameters.

Invariance of CTT and IRT item parameters across different participant groups

The degree of the CTT and IRT item parameters invariance was assessed by comparing the item parameter estimates within each measurement framework across two or more different participant groups: (1) lower and upper grade, (2) male and female, (3) Hungary, Latvia, Netherlands, Norway, Scotland, and Slovenia. Comparison included correlations between item parameter estimates in different groups (grade, gender, country) and comparing means of item parameters (paired t -test and ANOVA).

Invariance of CTT and IRT person and item parameters across different item sets

After both the CTT and the IRT item parameters had been obtained for all items in both item pools, four new tests (two mathematical and two science tests) of 12 items were constructed. For the construction of new CTT tests highly discriminative items with approximately normally distributed item difficulties were selected. For the new IRT tests highly discriminative and informative items with item difficulty parameters across wide range of theta values were selected (the target test information function was high through all levels of latent variable). Since the

item pools were rather small, only psychometric information was considered when constructing new tests.

Item and person parameters in these four new tests were compared with the original parameters, which were obtained with all items in both item pools. Comparison included correlations between the *new* and the *original* parameter estimates within each measurement framework as well as comparing mean differences of person/item parameters (paired *t*-test).

Results and Discussion

IRT Model Fit Assessment

The main goal of our study was to investigate both the comparability of CTT and IRT item parameters and the invariance of CTT and IRT item parameters in different conditions. As we were especially interested in item discrimination parameters, which usually represent the most important criteria for item selection, but also appear to be the most unstable item parameters, the 1PL model that would only allow the inspection of item difficulty parameters was not of interest. We were also not interested in the 3PL model for two reasons: (1) the 3PL model with a unique lower asymptote for each item would lead to different meaning of item difficulty for each item in the test and to different meaning of item difficulty for the same item in different item sets, and this would bias the results of comparison of item parameters across different item sets; (2) the 3PL model with a common lower asymptote for all items was not reasonable because not all items in our item pools were multiple-choice items.

The assessment of the IRT model fit indicates that the 2PL model comparing to the 1PL model fits to the data significantly better for all tests (the differences in $-2 \log \text{likelihoods}$) and also that the 3PL model fits to the data better than the 2PL for two tests only. According to the χ^2 test empirical ICC's differ significantly from expected ICC's in 2PL model for only some items in the math item pool, but for more than half of the items in the science item pool (Table 4). Since the χ^2 test in *BILOG-MG* can only be computed for tests with more than 20 items, only graphical comparison of expected and empirical ICC's could be made for new tests; the comparison showed that practically all items in new tests fit the 2PL model.

Table 4 also presents the main results of the (uni)dimensionality analysis. We can conclude that the assumption of unidimensionality holds to a reasonable extent in all new tests and in the math item bank. In the case of science item bank, however, the assumption of unidimensionality is clearly violated, which is probably the reason for a poor fit of many science items to the unidimensional 2PL model according to the results of χ^2 test. The unidimensional 2PL model is obviously not fully adequate for the science item bank (a multidimensional model would be appropriate), but was nevertheless used for modelling these data as well. It should be emphasized at

Table 4. Analysis of unidimensionality, reliability and IRT model fit for 2PL model for both item pools and for the new tests.

	Math			Science		
	Item Pool	New CTT test	New IRT test	Item Pool	New CTT test	New IRT test
% common variance	37.9	42.4	42.6	33.4	35.8	35.5
% explained common variance	51.7	73.0	72.4	38.4	57.1	57.7
Alpha reliability	0.893	0.833	0.819	0.819	0.724	0.710
GLB to reliability	0.920	0.864	0.857	0.865	0.784	0.775
% of misfitting items (χ^2 test) ^a	29.7	N/A	N/A	51.3	N/A	N/A

Note. N/A = not available

^a χ^2 test cannot be computed for new tests which consist of less than 20 items.

this point that all results involving the science item bank consequently demonstrate invariance of item/person parameters in the case of fairly heterogeneous test and a rather poor IRT model fit.

Comparability of CTT and IRT person and item parameters

The first goal of our study was to assess the comparability of CTT and IRT person and item parameters. As can be seen from the results in Table 5, CTT and IRT person parameters correlate very highly in both item pools, indicating that very similar math/science achievement estimates would be obtained regardless of the measurement framework. However, the comparison of the distributions of the CTT and IRT person parameters shows that this indication is only partly true. Since both math and science item banks contain more easy items than hard items, the distribution

Table 5. Correlations between CTT and IRT person and item parameters.

	Math Item Pool	Science Item Pool
Person parameters	0.984	0.990
Item parameters		
Item difficulty	0.972	0.922
Item discrimination	0.935	0.831

Note. In order to make the correlations positive, CTT item difficulty parameters were reversed so that higher values relate to more difficult items. Math item pool consisted of 39 items; science item pool consisted of 37 items.

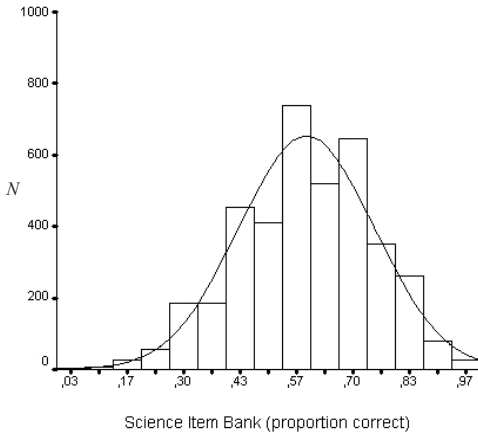


Figure 1. Distribution of proportion correct in science item pool. The curve represents normal distribution.

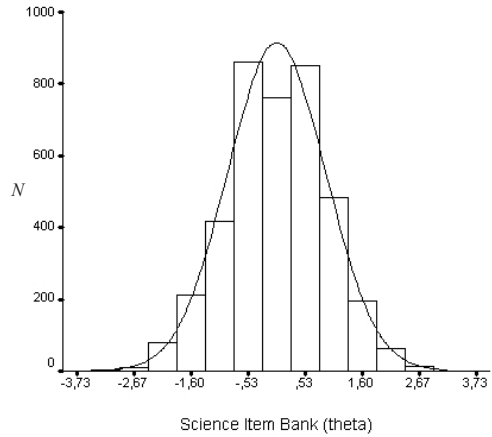


Figure 2. Distribution of theta values in science item pool. The curve represents normal distribution.

of the CTT person parameters in both item banks is negatively skewed. In contrast to this, the distributions of the IRT-based person parameters are more symmetrical, which could be the result of either item parameters being incorporated in the assessment of person parameters or the effect of prior normal distribution. The difference in person parameter distributions (for example, see Figures 1 and 2) could therefore be the result of the fact that the CTT based person parameters depend on the item parameters whereas the IRT person parameters do not, but, since neither true classical nor true IRT scores are known in this study, this conclusion would be unfounded.

The CTT and IRT item parameters correlate very highly, as well as the person parameters (also in Table 5). Item difficulty parameters correlate higher than item discrimination parameters, but the latter correlate very high as well in all conditions. Both the item difficulty and the item discrimination parameters are more comparable in the case of mathematical item pool, where the 2PL model fits the data better.

Invariance of CTT and IRT item parameters across different participant groups

The comparison of CTT and IRT item difficulty and item discrimination invariance across different groups of participants (regarding grade, gender and country) is presented in Table 6. Both CTT and IRT item difficulty parameters were estimated very consistently in all conditions. The highest consistence was found across the gender groups and the lowest one across different countries. The IRT item difficulties are slightly more consistently estimated than the classical item difficulties in most conditions for the math item pool, where the 2PL model fits to the data reasonably well, but less consistently in all conditions for the science item pool, where the model

Table 6. Comparison of IRT and CTT item parameters invariance across different groups of participants.

	Math Item Pool			Science Item Pool		
	Grade	Gender	Country	Grade	Gender	Country
Item difficulty						
CTT	0.946	0.979	0.787	0.983	0.983	0.859
IRT	0.953	0.980	0.783	0.973	0.979	0.835
Item discrimination						
CTT	0.894	0.945	0.657	0.770	0.903	0.601
IRT	0.930	0.943	0.694	0.698	0.877	0.514

Note. Correlations between item parameters estimated in two or more groups are presented. In case of country, single measure intraclass correlations (consistency) are shown. Math item pool consisted of 39 items, science item pool consisted of 37 items.

fit is rather poor. These results are coherent with presumption that IRT item parameter invariance can only be expected when the model fits the data (i.e., when the assumptions about the data are valid; Embretson & Reise, 2000; Hambleton, Swaminathan & Rogers, 1991). Considering that the IRT item parameter estimates for science data are not optimal within the unidimensional 2PL model, item difficulties are actually surprisingly invariant across different participant groups.

Both the CTT and the IRT item discrimination parameters are less invariant across different participant groups than the item difficulty parameters. The difference between item difficulty invariance and item discrimination invariance is especially high in the case of science data and for the IRT parameters; it seems that poor model fit has greater impact on the IRT item discrimination estimates than on the IRT item difficulties estimates. Similarly to item difficulty parameters, the consistency of IRT discrimination estimates is higher, compared to classical parameters, for most conditions in the case of math item pool, but lower in all conditions in the case of science item pool.

Another important issue remains to be discussed. The fact that item parameters were consistently estimated in two or more different participant groups does not mean that the item parameters are actually invariant. High correlations between item difficulties indicate that items which are easier, for instance, for students in the lower grade, also tend to be easier for students in the upper grade. The absolute values of item difficulty parameter estimates in the lower and in the upper grade however differ (see Table 8 and Figures 3 and 4 for example). Since students from upper grade probably know more mathematics and science than students from lower grade, this difference is expected for the CTT item difficulties but not for the IRT item difficulties. It should be noted that, since IRT person/item parameter scale is arbitrary, IRT item parameters obtained from different groups of participants should be rescaled before the comparisons between them can be made (Embretson & Re-

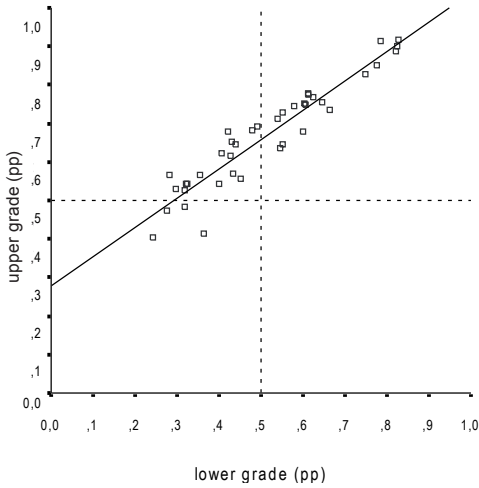


Figure 3. CTT item difficulty parameter estimates (pp) from math item pool in lower and upper grade.

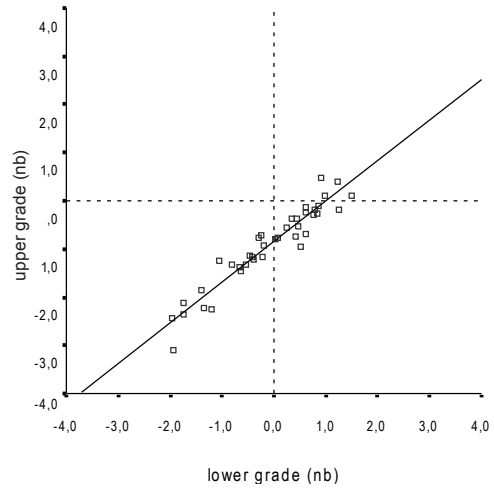


Figure 4. Non-rescaled IRT item difficulty parameter estimates (nb) from math item pool in lower and upper grade.

ise, 2000; Hambleton et al, 1991). But in Figure 4 non-rescaled item difficulties are shown; if these parameter estimates are rescaled, the IRT item difficulties are not only consistent but also (almost) invariant in the absolute sense. For the CTT item parameters, on the other hand, no such rescaling procedure exists.

Invariance of CTT and IRT person and item parameters across different item sets

After the item parameters for all items in math/science item pool had been obtained, two new tests were constructed from each item pool, the new CTT and the new IRT math/science test, respectively. It became evident, when selecting items for new tests, that our item pools were too small and did not have enough highly discriminating items. Consequently new CTT and IRT tests consist of up to three quarters of identical items; although – by design – different item selection criteria were used. This outcome, however, would not be a very likely one if item pools were large enough and of a high quality. The results regarding all four new tests are presented in the sequel but, since new CTT and IRT tests are very similar, the difference between them is not further discussed.

It should also be noted that item content was neglected when selecting items from item pools which resulted in rather different contents and dimensionality of the new tests comparing to the original item pools (see Table 4 for the results of dimensionality analysis). This is particularly true in the case of science data: new science tests contain items from some science achievement fields only and are con-

sequently closer to unidimensional, compared to all items from the pool, which are more diverse in contents.

In theory, IRT item parameters are invariant, not only across different participant groups, but also across different item sets, measuring the same latent variable. Accordingly, the IRT person parameter estimates should be very similar regardless of the item set they were obtained from. The property of item/person parameters invariance has implications for an important IRT application, i.e. computerized adaptive testing, where items are presented in different item sets and different participants usually reply to different items. Classical item discrimination parameters and person parameters, on the other hand, are clearly related to items in item set they are assessed from: the classical true score can only be defined with regard to a particular test; since test scores depend on the properties of the particular items in the test, so do the item-total correlations.

Table 7. Comparison of IRT- and CTT-based item and person parameters invariance across different item sets.

	New Math Tests		New Science Tests	
	CTT	IRT	CTT	IRT
CTT person parameters	0.942	0.908	0.870	0.857
IRT person parameters	0.943	0.921	0.883	0.869
Item difficulty parameters				
CTT	same	same	same	same
IRT	1.000	0.997	0.997	0.998
Item discrimination parameters				
CTT	0.975	0.914	0.631	0.700
IRT	0.975	0.885	0.408	0.377

Note. Correlations between the *original* and *new* person/item parameters are presented.

Person parameter invariance in different item sets was assessed by correlating the *original* (from item pools) and the *new* CTT and IRT person parameters and by comparing means of achievement estimates from item pools and new tests. It should be noted that correlations, presented in Table 7, only partially reflect the property of invariance as discussed above as item pools and new tests are not entirely different: all items from new tests are also included in item pools, what had influence on correlations, especially on the correlations between the *original* and *new* classical test scores. Furthermore, item pools contain more items than new tests and the person parameters were assessed from one measurement only (the context in which items appeared was the same). Although, due to this bias, classical test scores appear to be more invariant than they in fact are, IRT person parameters are slightly more

invariant when estimated from different item sets than CTT person parameters in all conditions. Correlations between *original* and *new* parameters are, again, higher in the case of mathematical items, because both new math tests measure achievement that is very similar to achievement measured in math item pool (all tests are relatively unidimensional, items are similar in contents), whereas the contents of new science tests is somewhat different comparing to contents of all items in science item bank (also new science tests are unidimensional, but science item pool is rather multidimensional). New science tests obviously do not measure *the same* science achievement (i.e., the same latent variable) as is measured by all items in science item bank; the degree of invariance of person parameters is consequently lesser. Considering this, the correlations between original and new IRT person parameters in the case of science achievement are actually surprisingly high. The results of comparing mean differences in person parameters obtained from item pools and new tests (presented in Table 9) confirm the results on person parameter invariance discussed above. They suggest that very similar person parameter estimates are obtained from item pools and new tests, since mean differences in parameters are very small, and also that IRT person parameters are more invariant when estimated from different item sets in all conditions. It is, again, possible that these high correlations and small differences between original and new IRT person parameters are, above all, the consequence of prior normal distribution used in parameter estimations; within our research design however this remains only speculation.

As previously discussed, IRT item parameters should be invariant in different item sets since they are, in theory, independent of all other items in the test. In CTT only item difficulties are independent of other items in the test, whereas item discrimination always depends on all other items. The comparison of item parameters invariance across different item sets is presented in Table 7. Correlations between *original* and *new* IRT item difficulty parameters are higher than correlations between item discrimination parameters. Very consistent estimates of IRT item difficulty parameters are obtained in new tests and both item pools, despite somewhat different contents of new (especially science) tests and the fact that item difficulty estimates were not optimal in the science item pool. The mean differences in IRT item difficulties across different item sets are also very small in all conditions (see Table 9). Classical item difficulty parameters are the same, as the same data set was used for estimation of *original* and *new* parameters.

The item discrimination parameters are, in comparison to the item difficulties, less invariant in all conditions; furthermore, the classical discrimination parameters are at least as much invariant as the IRT item discriminations (Table 7). Note that classical item discrimination parameters are unjustifiably high since items in item pools and new tests are not entirely different (in comparison of the *original* and the *new* CTT item discriminations the correlations between partly same items were calculated). Both the CTT and the IRT item discrimination parameters correlate lower in the case of science items, which is probably due to the difference in contents of

science item pool and new science tests: new science tests measure different and more unidimensional science achievement than all items in science item bank and items are simply not equally indicative for the *old* and the *new* science achievement. Apart from the content differences for science tests, poor invariance of the IRT item discrimination estimates, compared to the classical estimates, is probably also the result of suboptimal *original* item discrimination estimates. Apparently poor model fit for items in science item bank has the same effect on the IRT item discrimination invariance in different item sets than on the IRT item discrimination invariance across different participant groups. Despite the lower correlations of the item discrimination parameters as compared to the item difficulty parameters, the mean differences in item discrimination parameters are in general not significantly different when estimated from different item sets (Table 9) and although the differences between the *original* and *new* IRT item discrimination parameters for science data are generally larger than the differences for math data, these differences are also statistically insignificant, because the standard errors of means are higher in this case as well.

Conclusions

The present study attempted to address the questions of comparability of CTT and IRT person/item parameters, invariance of CTT and IRT item parameters across different groups of participants and invariance of CTT and IRT person/item parameters across different item sets.

Our main conclusions regarding the first question are that the IRT and the CTT person parameters are highly comparable and also that item difficulties and item discriminations are very comparable, with the latter showing somewhat lower comparability. Similar results regarding IRT and CTT parameters comparability were also found by Courville (2005), Fan (1998) and MacDonald and Paunonen (2002).

As for our second research question, the results of our study also support previous findings by Fan (1998) and MacDonald and Paunonen (2002) that CTT and IRT item parameters show very similar invariance property across different participant groups, with item difficulty parameters being more invariant than item discrimination parameters. But, in addition to this, the results of this study also show that IRT item parameters are generally more invariant than CTT parameters in the case of good IRT model fit, whereas CTT item parameters are more invariant in the case of poor IRT model fit. This conclusion is very important for application of IRT in test constructing as it implies (1) that the question of IRT model fit should be addressed very carefully when using IRT for item calibration, and (2) that, if an appropriate IRT model is used for modelling item responses, theoretically superior IRT item parameters would also be empirically superior and could therefore be used for item calibration on non-representative groups of participants. It should be noted here that,

although CTT item parameters also proved to be very consistently estimated in different groups of participants, the absolute values of item parameters differ and that, in contrary to the IRT item parameter estimates, there is no rescaling procedure, which would place CTT item parameter estimates from different participant groups on a common scale (Embretson & Reise, 2000; Hambleton et al, 1991).

The invariance property of person and item parameters across different item sets, an object of our third research question, also has an important consequence in test application: if person and item parameters are invariant across different item sets, different items can be used for measuring the same latent variable. Since the same data set was used for parameter estimation in item pools and new tests and, consequently, CTT item difficulties stayed the same, it was not possible to compare the invariance property of item difficulty parameters in both measurement frameworks; however, the IRT difficulty parameters have proven to be very invariant across different item sets. In contrary to theoretical and intuitive expectations, the results of this study also show that the CTT item discrimination parameters are just as or more invariant across different item sets than the IRT discrimination parameters and, as expected, that IRT item discriminations are invariant only if the model fits the data. We also found invariant person parameter estimates across different item sets, with the IRT person parameters being more invariant in all conditions. This is very interesting with regard to the IRT person parameters related to the science data, since the test with all items from the item pool and the new tests are obviously not measuring the same latent trait and since the IRT discrimination parameters are not invariant across item sets, and is therefore possibly only the effect of prior normal distribution used in the MML estimation method for IRT person parameters.

Of course, the present study, as it is an empirical study, has its share of limitations. First of all, item pools were limited in both item difficulties and item discriminations. The characteristics of the particular items that were used in the study could have some effect on the comparability of CTT and IRT person/item parameters and on the person/item parameters invariance properties. Larger item pools, with items varying more in item difficulty and item discrimination, or simulation studies that would manipulate different item characteristics should be used in future studies to determine whether the limited item pool also limits the generalization of the results from this study. Limited item pools were also related to item selection for new tests. An important aim at the beginning of this study was to determine whether the test construction method influences the properties of the resulting test. Unfortunately, because item pools were limited in the number of highly discriminative and informative items, very similar new tests were constructed using both CTT and IRT item selection criteria and possible differences in test construction could not be revealed. The second obvious shortcoming of this study was that the true person and item parameters were unknown. Because of this, only the comparison between CTT and IRT item/person parameters could be made and, for instance, the question whether the difference in distributions of CTT and IRT person parameters for item pools is

the result of the effect of the easiness of the test on CTT person parameters or the effect of prior normal distribution on IRT person parameters, could not be answered. The information truly valuable to a psychometric practitioner in this case would, of course, not be that of the comparability of IRT and CTT parameters, but about accuracy of parameter estimates. Again, we propose further investigation, possibly a simulation study that would overcome this limitation.

The fact that majority of the items allowed for guessing was not taken into account explicitly in our analyses. This should not be a source of bias against either paradigm. On one hand, guessing degrades the fit of the IRT models, because the empirical ICC can not be expected to approach zero when the theta value is small. On the other hand, guessing degrades the CTT-based measurement, too, because it narrows the range of item difficulties and because the difficulty parameters are not directly comparable across items with different probability of a random success.

Overall, the findings from this study indicate that CTT and IRT item/person parameters are very comparable and show similar invariance properties when estimated across different participant groups or across different item sets. Despite the high similarity between CTT and IRT parameters, the results also demonstrate that theoretical advantages (invariance property) of the IRT parameters are, at least to some extent, reflected in their empirical superiority, but only if an appropriate IRT model is used for modelling the data, the condition that is difficult to satisfy and even to assess appropriately in practice. Our findings also demonstrate that, with the MML method used for the IRT parameter estimation, model misfit influenced the item parameters invariance but not the person parameters invariance, what suggests that prior normal distribution has a great impact on person parameter estimates. Further investigation, preferably with simulated data with known true values of item/person parameters, is needed to provide missing guidelines about the assessment of IRT model fit, about the extent of model misfit that would still yield sufficiently invariant item and person parameters and to determine the influence of different IRT estimation methods on the accuracy of IRT person/item parameter estimates.

References

- Courville, T. G. (2005). *An empirical comparison of item response theory and classical test theory item/person statistics*. Unpublished doctoral dissertation, Texas A&M University. Retrieved November 5, 2007 from <http://txspace.tamu.edu/bitstream/handle/1969.1/1064/etd-tamu-2004B-EPSY-Courville-2.pdf?sequence=1>.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Fan, X. (1998). Item response theory and classical test theory: An empirical comparison of their item/person parameters. *Educational and Psychological Measurement*, 58, 357–381.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item re-*

- sponse theory*. Newbury Park, CA: Sage.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.
- MacDonald, P., & Paunonen, S. V. (2002). A Monte Carlo comparison of item and person parameters based on item response theory versus classical test theory. *Educational and Psychological Measurement, 62*, 921–943.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- Ten Berge, J. M. F., & Kiers, H. A. L. (1991). A numerical approach to the approximate and the exact minimum rank of a covariance matrix. *Psychometrika, 56*, 309–315.
- Ten Berge, J. M. F., Snijders, T. A. B., & Zegers, F. E. (1981). Computational aspects of the greatest lower bound to the reliability and constrained minimum trace factor analysis. *Psychometrika, 46*, 201–213.
- Ten Berge, J. M. F., & Sočan, G. (2004). The greatest lower bound to the reliability of a test and the hypothesis of unidimensionality. *Psychometrika, 69*, 613–625.
- Zimowski, M. F., Muraki, E., Mislevy, R. J., & Bock, R. D. (1996). *BILOG-MG: Multiple-group IRT analysis and test maintenance for binary items*. Chicago: Scientific Software.

Table 8. Comparison of the IRT- and the CTT- based item parameters invariance across different groups of participants (results of *t*-test for grade and gender; ANOVA for country).

Item difficulty	Grade				Gender				Country					
	<i>d</i>	<i>df</i>	<i>t</i>	<i>p</i>	<i>d</i>	<i>df</i>	<i>t</i>	<i>p</i>	<i>df_i</i>	<i>df_j</i>	<i>F</i>	<i>p</i>	η^2	
Math Item Bank	CTT	-0.156	38	-16.22	0.000	-0.002	38	-0.38	0.636	5.0	190.0	25.06	0.000	0.397
	IRT	0.830	38	17.68	0.000	0.008	38	0.30	0.762	3.9	146.2	22.84	0.000	0.375
Science Item Bank	CTT	-0.096	36	-13.27	0.000	-0.029	36	-4.18	0.000	3.8	135.6	2.16	0.081	0.057
	IRT	0.752	36	11.88	0.000	0.172	36	3.01	0.005	5.0	180.0	3.56	0.004	0.090
Item discrimination														
Math Item Bank	CTT	-0.019	38	-2.81	0.008	0.009	38	1.82	0.076	3.7	139.4	7.31	0.000	0.161
	IRT	-0.048	38	-3.38	0.002	0.013	38	1.06	0.294	3.5	132.1	5.52	0.001	0.127
Science Item Bank	CTT	0.020	36	2.34	0.025	-0.019	36	-3.32	0.002	5.0	180.0	13.65	0.000	0.275
	IRT	0.015	36	0.73	0.468	-0.032	36	-2.48	0.018	4.1	147.2	6.08	0.000	0.145

Note. *d* = Cohen's effect size measure; η^2 = proportion of variance explained by factor.

Appendix

Table 9. Comparison of IRT- and CTT-based person/item parameters invariance across different item sets (results of paired *t*-test between the original and the new parameter estimates).

	New CTT constructed tests			New IRT constructed tests		
	<i>d</i>	<i>t</i>	<i>p</i>	<i>d</i>	<i>t</i>	<i>p</i>
Person parameters						
New Math Tests						
CTT	0.093	3947	46.50	0.000	3947	-1.56
IRT	0.001	3947	0.22	0.824	3947	0.33
New Science Tests						
CTT	0.090	3950	45.36	0.000	3950	29.68
IRT	-0.001	3950	-0.16	0.873	3950	-0.12
Item difficulty						
New Math Tests						
CTT	same	-	-	-	-	-
IRT	0.008	11	3.12	0.010	11	-0.06
New Science Tests						
CTT	same	-	-	-	-	-
IRT	-0.013	11	-0.55	0.597	11	-0.29
Item discrimination						
New Math Tests						
CTT	-0.013	11	-2.86	0.016	11	-0.99
IRT	0.018	11	0.90	0.387	11	1.03
New Science Tests						
CTT	-0.010	11	-0.76	0.469	11	-0.87
IRT	0.038	11	0.88	0.397	11	0.67

Note. *d* = Cohen's effect size measure.

Metaanaliza in njen pomen za psihološko metodologijo

*Andrej Kastrin**

Univerzitetni klinični center Ljubljana, Inštitut za medicinsko genetiko, Ljubljana

Povzetek: Metaanaliza je posebna oblika statistične analize, v kateri na sistematičen način združujemo rezultate posameznih med seboj neodvisnih študij. Glavni namen prispevka je opredeliti metodo metaanalize, opozoriti na nekatere značilnosti in posebnosti pri uporabi ter spodbuditi raziskovalce k njeni uporabi pri raziskovalnem delu. Pravilno izvedena metaanaliza omogoča sistematično spremljanje najnovejših znanstvenih spoznanj, učinkovitejšo izrabo obstoječih podatkov, pomembno prispeva h kakovosti obstoječega znanja o določenem proučevanem fenomenu in lahko služi kot podpora pri gradnji novih raziskovalnih domnev. Zamisel o možnosti združevanja podatkov in rezultatov neodvisnih študij je stara že skoraj pol tisočletja, temelje moderne metaanalize pa je pred dobrimi tremi desetletji postavil Glass, ko je na metodološko rigorozen način ovrgel Eysenckovo domnevo o ničnosti učinka psihoterapije. V članku osvetlimo kronološki potek razvoja metaanalize, predstavimo splošen potek metode ter njene glavne statistične koncepte. Podrobneje predstavimo pojem velikosti učinka, problematiko heterogenosti študij ter dva modela združevanja rezultatov študij v skupno oceno velikosti učinka: model stalnih in model slučajnih učinkov. Na primeru podamo zgled za vizualizacijo rezultatov metaanalize s pomočjo drevesnega in lijakastega diagrama. Z namenom enostavne in hitre podpore k metaanalitični metodologiji smo razvili spletni strežnik RMetaWeb, ki raziskovalcem ponuja možnost interaktivne analize podatkov. Gre za prvo tovrstno spletno orodje, katerega jedro predstavlja okolje R za statistično analizo in grafiko.

Ključne besede: psihološka metodologija, statistika, metaanaliza

Meta-analysis: Its role in psychological methodology

Andrej Kastrin

University Medical Centre Ljubljana, Institute of Medical Genetics

Abstract: Meta-analysis refers to the statistical analysis of a large collection of independent observations for the purpose of integrating results. The main objectives of this article are to define meta-analysis as a method of data integration, to draw attention to some particularities of its use, and to encourage researchers to use meta-analysis in their work. The benefits of meta-analysis include more effective exploitation of existing data from independent sources and contribution to more powerful domain knowledge. It may also serve as a support tool to generate new research hypothesis. The idea of combining results of independent studies addressing the same research question dates back to sixteenth century. Meta-analysis was reinvented in 1976 by Glass, to refute the conclusion of an eminent colleague, Eysenck, that psychotherapy was essentially ineffective. We review some major historical landmarks of meta-analysis and its statistical background. We present the concept of effect size measure, the problem of heterogeneity and two models which are used to combine individual effect sizes (fixed and random

* Naslov/Address: asist. Andrej Kastrin, univ. dipl. psih., Univerzitetni klinični center Ljubljana, Inštitut za medicinsko genetiko, Štajmerjeva ulica 3, 1000 Ljubljana, Slovenija, e-pošta: andrej.kastrin@guest.arnes.si

effect model) in great details. Two visualization techniques, forest and funnel plot graphics are demonstrated. We developed RMetaWeb, simple and fast web server application to conduct meta-analysis online. RMetaWeb is the first web meta-analysis application and is completely based on R software environment for statistical computing and graphics.

Key words: psychological methodology, statistics, meta-analysis

CC = 2200

Kot odgovor na izzive zajemanja, shranjevanja in upravljanja z velikimi količinami podatkov, informacij in znanja se je v zadnjem desetletju uveljavilo raziskovalno področje, ki se imenuje odkrivanje zakonitosti iz podatkov. Gre za odkrivanje implicitnih, doslej neznanih in potencialno uporabnih zakonitosti iz podatkov, z namenom učinkovitejšega odločanja, razvrščanja in napovedovanja. Danes, ko je statistika ena od vodilnih znanstvenih disciplin in eden od paradnih konjev področja odkrivanja zakonitosti iz podatkov, lahko psihologija s ponosom ugotovi, da je mnogo pomembnih konceptov, modelov in teorij v zakladnico statističnega znanja prispevala prav sama. Najbolj pomembni prispevki so Stevensova tipologija merskih lestvic, metoda faktorске analize s konceptom latentnih spremenljivk, multidimenzionalno skaliranje ter široko področje testne teorije. Tudi mnogo pomembnih statistikov je po svoji osnovni izobrazbi psihologov, npr. Donald B. Rubin, Leland Wilkinson, Paul Murrell, če omenimo le nekatere. Če že ne najpomembnejši, pa zagotovo najbolj odmeven metodološki prispevek, vsaj sodeč po številu objav v mednarodnih bibliografskih zbirkah, pa je psihologija dala na področju metaanalize.

Popularna Wikipedia takole povzema razumevanje pojma metaanaliza: "Metaanaliza je statistična metoda, namenjena združevanju rezultatov večjega števila študij, ki se ukvarjajo s proučevanjem podobnega raziskovalnega problema." Metaanalizi nadreden koncept je sistematičen pregled literature (Torgerson, 2003). Gre za metodo pregleda literature, povzemanja in zbiranja kvalitativnih dokazov o nekem raziskovalnem problemu. Metaanalizo opredelimo bolj specifično, kot tehniko pregleda literature, z natančno določeno metodologijo in kvantifikacijo rezultatov podobnih študij s standardno metriko, ki omogoča uporabo statističnih metod kot sredstva analize (Rosenthal, 1991; Wachter, 1988; Wolf, 1986). Pri metaanalizi se bibliografski viri ne uporabljajo za definicijo raziskovalnega problema, ampak njihov pregled predstavlja samostojen problem, ki privede do teoretičnih in empiričnih zaključkov, ki lahko spremenijo ali dopolnijo znanje na nekem področju znanstvenega proučevanja. Njene korenine segajo na področje psihologije in pedagogike, kasneje pa se je močno razširila v praktično vse temeljne in aplikativne znanstvene vede. Metaanaliza, skupaj z nepogrešljivo faktorško analizo in čedalje bolj uporabljeno metodologijo strukturnih enačb, predstavlja močen paket podpore modernemu raziskovalnemu delu v psihologiji.

Zakaj metaanaliza

Povezovanje podatkov in informacij, nenehen razvoj novega znanja in njegovo plemenitenje v praksi morda še nikoli niso bili tako pomembni kot danes. Zlasti v svetu zunaj laboratorijev, inštitutov in univerz pomeni golo kopičenje znanja, brez njegovega pretapljanja v rast in razvoj, izgubo konkurenčnosti.

Kopičenju znanja na nekem znanstvenem področju lahko sledimo z dvema komponentama: zadostnostjo in stabilnostjo (Schmidt, 1992). Komponenta zadostnosti se nanaša na vprašanje količine študij, ki jih potrebujemo, da zadovoljivo opišemo nek fenomen oz. problemsko domeno, komponenta stabilnosti pa na vprašanje skladnosti obstoječega znanja z znanjem, ki ga dobimo na osnovi novih raziskovalnih izsledkov pri ponovljenih merjenjih istega fenomena. Medtem ko na zadostnost vpliva predvsem integracija raziskovalčeve ustvarjalnosti na eni ter zadovoljevanje kriterijev znanstvene uspešnosti na drugi strani, je skladnost bolj objektivna in lažje preverljiva kategorija. V ožjem pomenu besede jo lahko skrcimo na zanesljivost merjenja določenega predmeta znanstvenega proučevanja.

Z vprašanjem zanesljivosti se srečujemo v vseh znanstvenih disciplinah, ki poskušajo svoje raziskovalne domneve preveriti z empiričnimi izsledki. Zanesljivi rezultati so ključnega pomena za doseganje osnovnega cilja znanstvenega raziskovanja, tj. ugotavljanja zakonitosti, ki nam omogočajo pojasnjevanje in napovedovanje opazovanih pojavov (Ferligoj, Leskošek in Kogovšek, 1995). Zanesljivost v širšem smislu pomeni, da bomo s ponavljanjem meritev istega pojava v enakih ali vsaj primerljivih okoliščinah dobili primerljive rezultate. Zanesljivost merjenja je tem večja, čim bolj so razlike v izmerjenih vrednostih posledica dejanskih sprememb merjenega pojava in čim manjši je vpliv slučajnih dejavnikov. V psihologiji je najbolj pereč problem, povezan z zanesljivostjo merjenja, vprašanje relativno majhnih vzorcev. Sedlmeier in Gigerenzer (1989) sta pred leti poročala, da je povprečna statistična moč¹ zaključevanja objavljenih študij v eni od psiholoških revij znašala komaj dobrih 40 %. S podobnim problemom se soočajo praktično vse znanstvene discipline, katerih raziskave so zasnovane na teoriji vzorčnega zaključevanja (prim. npr. Balding, 2006)

Smiseln odgovor na problem zanesljivosti posameznih študij, kliničnih poskusov oz. eksperimentov ponuja njihova integracija v obliki metaanalize. Metaanaliza omogoča večjo moč statističnega zaključevanja pri opazovanju določenega fenomena ter natančno oceno njegove variabilnosti (raztrosa) med študijami. Ideja o združevanju podatkov večih med seboj neodvisnih študij je stara že dobrih 400 let, temelje moderne metaanalize pa je pred slabimi tremi desetletji postavil Gene V. Glass (Smith in Glass, 1977), ko je na metodološko rigorozen način ovrgel smelo

¹ Moč statističnega testa se nanaša na verjetnost zavrnitve ničelne hipoteze, ko ta dejansko ne drži (Cohen, 1992). Statistična moč nekega testa je odvisna od treh parametrov: (i) izbrane ravni tveganja α , (ii) željene velikosti učinka ter (iii) velikosti vzorca.

Eysenckovo tezo o ničnosti učinka psihoterapije. Ustrezno izvedena metaanaliza (i) ponuja sistematične, hitre in zanesljive odgovore na raziskovalne domneve, (ii) zaradi večje količine podatkov povečuje moč statističnega zaključevanja, (iii) daje pregled nad metodologijo izvedbe posameznih poskusov ter nenazadnje (iv) omogoča velik prihranek sredstev na račun ponovitvenih poskusov.

Metaanaliza skozi čas

Metaanaliza ima dolgo preteklost, a razmeroma kratko zgodovino. Metodo ponovljenih merjenj pri merjenju istega pojava je v znanost vpeljal danski astronom Tycho Ottesen Brahe konec 16. stoletja (Plackett, 1958). Kepler je svoje tri slavne zakone, ki opisujejo gibanje planetov okoli Zemlje, osnoval ravno na osnovi njegovih dolgoletnih meritev. Brahe je bil prvi, ki je za zmanjševanje sistematične napake pri merjenju uporabil matematični koncept aritmetične sredine, ki se je v znanosti utrdil šele dobro stoletje kasneje. Drug pomemben miselni preskok v teoriji merjenja, neposredno povezan z razvojem metaanalize, je kombinacija meritev različnih opazovalcev, ki jo je vpeljal francoski matematik in astronom Pierre-Louis Moreau de Mauperuis (Plackett, 1958). Pri merjenju dolžin poldnevniške (meridianske) stopinje si je pomagal z večimi neodvisnimi opazovalci, meritve povprečil in tako empirično potrdil pravilnost Newtonove teorije o sploščenosti Zemlje. V veliki meri so bili prav astronomi tisti, ki so postavili temeljne kamne sodobni teoriji merjenja. George Biddell Airy je leta 1861 ugotovitve svojih stanovskih kolegov povzel v znanstveni monografiji z naslovom "On the algebraical and numerical theory of errors of observations and the combination of observations". Prvi resen poskus združevanja kliničnih rezultatov je na začetku 20. stoletja izvedel Karl Pearson z združitvijo podatkov različnih študij, ki so proučevale vpliv cepiva proti tifoidni mrzlici na različnih vzorcih angleških vojakov (Pearson, 1904). Medicina je potrebovala skoraj 50 let, da je odkrila Pearsonov prispevek (Olkin, 1995). Drug pomemben oče metaanalize je bil slavni britanski statistik in Darwinov naslednik Ronald A. Fisher. V eni od sklepnih monografij (Fisher, 1970) je takole povzel bistvo svojega pogleda na problem integracije rezultatov različnih neodvisnih študij pri merjenju istega pojava:

Pri testiranju statistične značilnosti večih neodvisnih testov se včasih zgodi, da malo oz. noben test ni posamezno statistično značilen, združeni pa dajo vtis, da so verjetnosti (zavrnitve ničelne hipoteze, op. a.) nižje, kot bi bile dobljene po naključju. (str. 99)

Za razliko od Pearsona, ki je združil surove korelacijske koeficiente posameznih študij, sta Fisher in Leonard Tippett naredila korak dlje ter neodvisno drug od drugega izpeljala inovativen postopek združevanja p -vrednosti pri testiranju več neodvisnih ničelnih hipotez (Rosenthal, 1991; Tippett, 1931). Medtem ko je Tippettov prispevek

utonil v zakladnico statistične zgodovine, se Fisherjev obrazec uporablja še danes. Fisher je pokazal, kako lahko m neodvisnih p -vrednosti združimo v enotno mero statistične značilnosti, ki se porazdeljuje po χ^2 porazdelitvi z $2m$ stopnjami prostosti (Fisher, 1970):

$$\chi_{2m}^2 = -2 \sum_{i=1}^m \log_e(p_i)$$

Eden od bolj slavnih primerov uporabe te enačbe v praksi je Gordonov (Gordon, Loveland in Cureton, 1952) poskus združevanja rezultatov študij inteligentnosti, ki sta jih opravila McNemar in Terman (1936).

Okno v svet je metaanalizi odprl Glass. Bolj kot ne zaradi pozitivne osebne izkušnje z lastno psihoterapijo se je spustil v ostro polemiko z eminentnim Eysenckom, zlasti z njegovo trditvijo o ničnosti učinka psihoterapije (Eysenck, 1952, 1965). Glassa štejejo za utemeljitelja sodobne metaanalize, je avtor skovanke metaanaliza ter nosilec nove paradigme v razvoju znanosti (Glass, 1976; Smith in Glass, 1977). Pred dobrimi 30 leti je takole zapisal:

Metaanaliza se nanaša na analizo analize. S terminom označujem statistično analizo velike zbirke rezultatov posameznih študij z namenom integracije novih spoznanj. Predstavlja močno alternativno dosedanji vzročni in pripovedni razlagi rezultatov in lahko služi kot podpora pri osmišljanju velike količine raziskovalnih podatkov. (Glass, 1976, str. 3)

Metaanaliza torej ni le suhoparen skupek statističnih obrazcev, ampak dodelan metodološki okvir za izkop novega znanja iz podatkov in njihovo osmišljanje (Schmidt, 1992). Istega leta je Robert Rosenthal (1976) objavil knjigo z naslovom "Experimenter effects in behavioral research", v kateri je predstavil koncept mer velikosti učinka in s tem sprožil močno kritiko klasične uporabe statističnih testov. Z uvedbo od velikosti vzorca neodvisnih mer razlik med rezultati merjenih spremenljivk je bilo tako dostopno tudi močno statistično orodje za primerjanje različnih študij med seboj. Eden najpomembnejših avtorjev s področja mer velikosti učinka je Jacob Cohen, ki je temelje kritike klasičnega testiranja statističnih domnev predstavil v članku s pomenljivim naslovom "The Earth is round ($p < .05$)" (Cohen, 1994). Leta 1977 je Glass skupaj s sodelavko Mary Lee Smith objavil članek, v katerem sta analizirala 375 neodvisnih študij s skupaj več kot 40.000 udeleženci, ki so z različnimi tehnikami in raziskovalnimi metodami proučevale učinke zdravljenja v različnih smereh psihoterapije, in Eysenckovo domnevo ovrgla (Smith in Glass, 1977). Eysenck je do konca svojega ustvarjalnega življenja ostal vnet nasprotnik takega pristopa k raziskovanju. Najbolj znan in največkrat citiran je njegov članek s provokativnim naslovom "Meta-analysis is an exercise in mega-silliness", v katerem je metaanalizo označil za nevredno metodo resnega znanstvenega dela (Eysenck, 1978).

Od vseh znanstvenih disciplin je metaanalizo najbolj unovčila biomedicina. Rečemo lahko, da je danes v biomedicini metaanaliza ena od najbolj uporabljenih metod zbiranja, analize in interpretacije raziskovalnih rezultatov (Egger, Ebrahim in Smith, 2002). Za razliko od drugih znanstvenih disciplin, je v biomedicini metaanaliza postala prava akademska industrija. Metaanaliza je postavila temelje t. i. znanstveno utemeljene medicine (angl. *evidence based medicine*). Znanstveno utemeljena medicina zajema postopek sistematičnega iskanja, ocenjevanja in nenazadnje tudi uporabe sodobnih raziskovalnih izsledkov kot temelj kliničnega odločanja ter pridobiva na popularnosti v številnih medicinskih disciplinah (Sackett, Rosenberg, Gray, Haynes in Richardson, 1996). Gre za razmeroma novo paradigmo, ki bo tudi medicini počasi omogočila uporabo pravega znanstvenega načina raziskovanja in njenega metodološkega aparata. Z namenom zagotavljanja kvalitetnih sistematičnih pregledov so po svetu ustanovili posebne raziskovalne centre, ki za potrebe medicine in njej sorodnih strok že več kot deset let opravljajo sistematične preglede in metaanalize. Najpomembnejši in najbolj organizirani so Cochranovi centri, poimenovani po angleškem epidemiologu Archiju Cochranu (Hill, 2000). Slovenija spada pod okrilje italijanske podružnice s sedežem v Milanu. Eden od najpomembnejših dosežkov združenja je Cochranova knjižnica (<http://www.cochrane.org/>), spletna zbirka sistematičnih pregledov in metaanaliz s področja biomedicine, med katerimi najdemo tudi precej zadetkov s področja psihologije in psihiatrije.

V zadnjem času so pomembne metodološke prispevke k metaanalizi dodali Nambury S. Raju, Larry V. Hedges, Ingram Olkin, John E. Hunter, Jacob Cohen, and Frank L. Schmidt. Z izjemo dveh, Hedgesa in Olkina, so (bili) vsi ostali avtorji eminentni raziskovalci na področju psihološke metodologije.

Metaanaliza je od prvih resnih poskusov pred 30 leti do danes postala močna veja teoretičnega, še bolj pa aplikativnega znanstvenega raziskovanja. V bibliografski zbirki PsycINFO je od leta 1865 do danes (10. 6. 2008) zbranih 1.168 zapisov, ki na kakršenkoli način omenjajo metaanalizo, od tega je 184 zapisov indeksiranih v specializirani zbirki PsycARTICLES. Googlov Učenjak (angl. *Google Scholar*), iskalnik po arhivih odprtega dostopa in spletnih mestih akademskih institucij nam ponudi 20.800 različnih zadetkov v povezavi z metaanalizo, medtem ko je v zbirki Medline² takih zadetkov kar 30.945.

Računsko jedro metaanalize

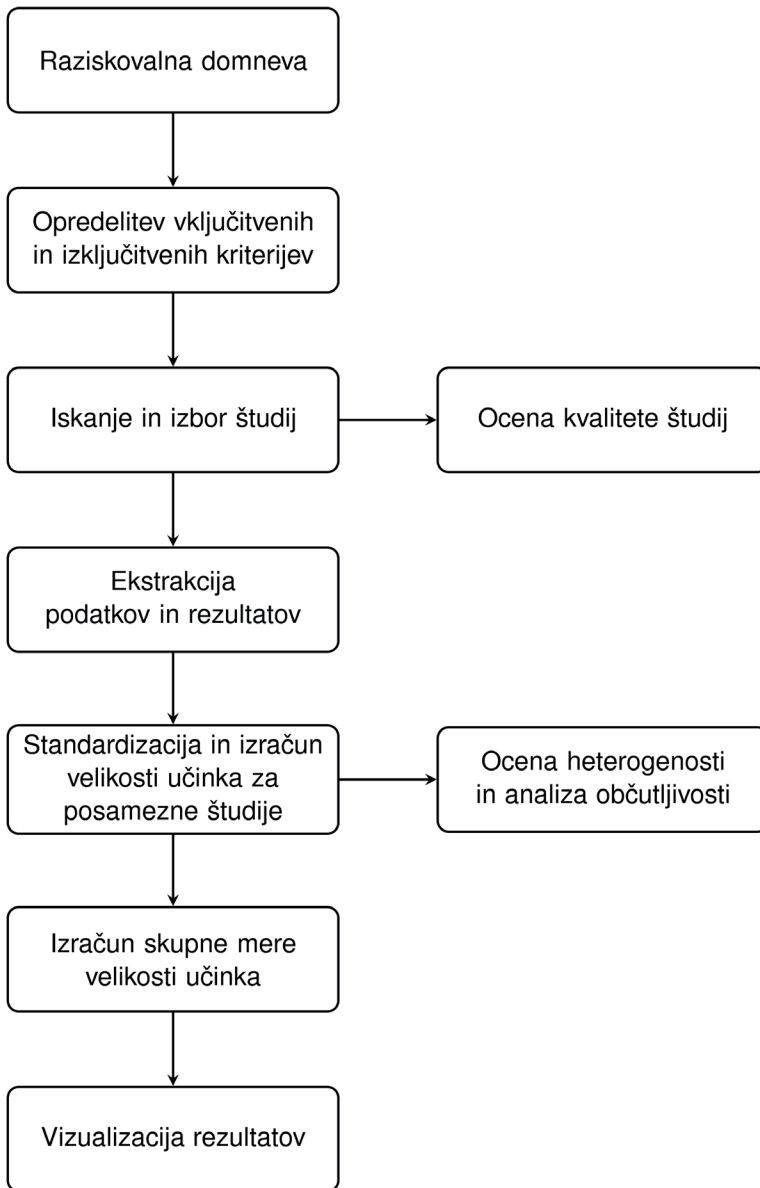
Prvi korak vsake raziskave je jasna opredelitev raziskovalnega problema. Tako kot pri vsaki eksperimentalni ali korelacijski raziskavi moramo tudi v primeru metaanalize natančno opredeliti ničelno domnevo. Jasno moramo opredeliti vključitvene

² Medline je bibliografska zbirka s področja biomedicine in znanosti o življenju in indeksira tudi precej revij s področja psihologije.

in izključitvene kriterije, ki nam bodo omogočali kasnejšo selekcijo med enotami, ki jih bomo vključili v našo raziskavo. Denimo, da želimo s pomočjo metaanalize ovrednotiti moč povezave med izraženostjo ekstravertnosti in potezo iskanja dražljajev v splošni populaciji. Primer vključitvenega kriterija bo v takem primeru npr. zajem vseh tistih študij, ki so kot merska pripomočka uporabili Eysenckov osebnostni vprašalnik EPQ-R za merjenje ekstravertnosti ter Zuckermanovo lestvico iskanja dražljajev SSS-VI, primer izključitvenega kriterija pa izpust študij, ki so testirale skupine udeležencev, za katere je znano, da imajo potezo iskanja dražljajev močnejše izraženo (npr. alpinisti, piloti, vojaki). Kriteriji lahko zajemajo tudi ostale parametre študij, npr. število udeležencev, tip eksperimentalne zasnove, leto in jezik objave študije itd. S pomočjo čimbolj učinkovitih iskalnih strategij nato iz vseh dostopnih bibliografskih in ostalih podatkovnih skladišč poiščemo vse zadetke, ki se nanašajo na predmet naše raziskave. V ožji izbor uvrstimo vse tiste zadetke, ki zadostijo vključitvenim kriterijem, ter izločimo zadetke, ki zadostijo izključitvenim kriterijem. Prvemu situ lahko dodamo še drugo sito, s katerim preverjamo kvaliteto posameznih študij. V ta namen so najbolj uporabni različni standardizirani postopki (npr. ocenjevalne liste), s katerimi ocenimo primernost posamezne študije (Wood, 2008). Faze ocenjevanja primernosti študij za vključitev v metaanalizo je smiselno izvesti paroma, z dvema neodvisnima ocenjevalcema. Sledi faza ekstrakcije podatkov, v kateri pripravimo podatkovno zbirko, ki za vsako študijo vključuje podatke o izmerjenih spremenljivkah ter za študijo specifične parametre. Slednji nam lahko v kasnejših korakih raziskovanja služijo kot sospremenljivke pri dodatnih analizah³. Surove podatke standardiziramo, s čimer dosežemo primerljivost rezultatov med študijami, ter izračunamo velikost učinka po posameznih študijah. S podporo namenskih računalniških programov nato izračunamo še oceno skupne velikosti učinka, katere stabilnost lahko potem preverjamo s sistematičnim izključevanjem posameznih študij (analiza občutljivosti). Povzetek opisanih korakov je prikazan na Sliki 1.

Glavni statistični koncept, s katerim operiramo v metaanalizi, je mera velikosti učinka (Hedges in Olkin, 1985). Dober teoretično-praktični uvod v problematiko mer velikosti učinka ponuja članek domačih avtorjev Cankarja in Bajca (2003). Mere velikosti učinka lahko taksonomsko razdelimo v dve veliki skupini. V prvo skupino uvrščamo mere standardiziranih razlik med proučevanimi skupinami, v drugo pa mere povezanosti. Nekateri avtorji tej grobi razvrstitvi dodajajo še intervale zaupanja (Hedges in Olkin, 1985). Pri standardiziranih razlikah gre običajno za opisovanje razdalj med aritmetičnimi sredinami proučevanih skupin v enotah standardnih odklonov, pri merah povezanosti pa za stopnjo povezanosti med učinkom in odvisno spremenljivko. Glede na tip spremenljivk je tako prva skupina mer uporabna zlasti na merskih ravneh, ki zadostijo vsaj ordinalni merski lestvici, druga skupina pa za

³ Klasični primer je npr. metaregresija. Odvisno spremenljivko predstavlja velikost učinka študije, prediktorske spremenljivke pa predstavljajo parametri študij (npr. starost, spol, socialno-ekonomski status udeležencev, tip eksperimentalnega načrta itd.).



Slika 1. Diagram poteka skozi glavne faze metaanalize.

nominalne (imenske) podatke. Med najbolj znane mere za ocenjevanje razlik spadajo Cohenov d , Glassov Δ in Hedgesova g statistika, med mere povezanosti pa statistike, ki računajo delež pojasnjene variance (npr. Pearsonov r koeficient korelacije, Cramerjev V koeficient povezanosti med nominalnima spremenljivkama itd.) ter ostale, v psihologiji manj znane mere povezanosti, izpeljane iz verjetnostnega računa (npr. razmerje obetov, relativno tveganje, pripisljivo tveganje itd.). Odličen pregled

različnih mer velikosti učinka ponuja Cohenova monografija (Cohen, 1988).

Heterogenost se nanaša na raztros posameznih ocen velikosti učinka preko študij (Whitehead, 2002). V grobem ločimo dve vrsti heterogenosti, vzorčno heterogenost in metodološko heterogenost. Prva se nanaša na razlike v vzorcih udeležencev, npr. razlike v spolu, starosti, socialno-ekonomskem položaju itd., druga pa na razlike v uporabljeni metodi merjenja proučevanega predmeta raziskave, na razlike v eksperimentalnih zasnovah in kvaliteti izvedbe posameznih študij. Oba vira heterogenosti imata za posledico statistično heterogenost, ki se nanaša na opazovane razlike med velikostmi učinkov posameznih študij. Določena stopnja klinične in metodološke heterogenosti med študijami vedno obstaja, v metaanalizi pa si prizadevamo, da je čim manjša.

Celotna teorija statističnega opisovanja in zaključevanja temelji na predpostavkah in modelih, zato moramo tudi pri metaanalizi opredeliti statistični model, na osnovi katerega bomo iz dosežkov po posameznih študijah sklepali na povprečno oceno velikosti učinka (Hedges in Olkin, 1985). V metaanalizi se najpogosteje uporabljata dva statistična modela (Hunter in Schmidt, 2000; Kisamore in Brannick, 2008): model stalnih učinkov (angl. *fixed effect model*) in model slučajnih učinkov (angl. *random effect model*). Model stalnih učinkov predpostavlja, da vključene študije ocenjujejo isti učinek oz. da so učinki posameznih študij vzorčne vrednosti iste populacije (Mantel in Haenszel, 1959). Na ta način upoštevamo le raztros znotraj posameznih študij. Predpostavka tega modela je torej, da vse vključene študije uporabljajo enako metodo merjenja, vzorci udeležencev pa so med seboj homogeni. Model slučajnih učinkov po drugi strani predpostavlja, da so vključene študije naključno vzorčene iz različnih populacij študij, ki imajo različne učinke (DerSimonian in Kacker, 2007; DerSimonian in Laird, 1986). Pri tem poleg raztrosa znotraj posameznih študij upoštevamo tudi raztros med študijami. Pri uporabi modelov moramo biti pazljivi, saj lahko modela na istih podatkih pripeljeta do popolnoma različnih rezultatov. V primeru, da so študije med seboj homogene, modela stalnih in slučajnih učinkov vrneta praktično primerljive rezultate. Teoretično gledano je prav heterogenost med študijami tista, ki definira izbiro statističnega modela. V primeru heterogenosti med študijami, ki ni posledica razlik nad populacijo vključenega univerzuma udeležencev v merjeni lastnosti, ampak je posledica raztrosa med posameznimi študijami, pa uporabimo model slučajnih učinkov. Za preverjanje predpostavke homogenosti posameznih študij lahko uporabimo različne testne statistike in grafične metode. Najbolj enostaven indikator statistične heterogenosti je slabo prekrivanje intervalov zaupanja, za numerično oceno pa uporabimo različne mere, izpeljane iz χ^2 in F preizkusa. Heterogenost študij sama po sebi ni nujno ovira. Resda je z vidika združevanja rezultatov moteča, vendar nam lahko služi tudi kot indikator vsebinskih razlik med študijami. Nekateri avtorji v tem okviru govorijo celo o dvojni naravi metaanalitičnih študij: njeni klasični, analitični obliki dodajajo še eksploratorni vidik, katerega namen je odkrivanje razlik in pojasnjevanje virov heterogenosti študij (Schmidt, 1992).

Običajno je tako, da podatkom najprej priredimo model s stalnimi učinki in preverimo njegovo prilaganje podatkom. Če empirični model statistično značilno

odstopa od pričakovanega, podatkom priredimo še model s slučajnimi učinki (Field, 2003). Model s slučajnimi učinki je v statističnem smislu bolj konzervativen in vrne širši interval zaupanja končne ocene velikosti učinka kot model s stalnimi učinki.

Vizualizacija rezultatov

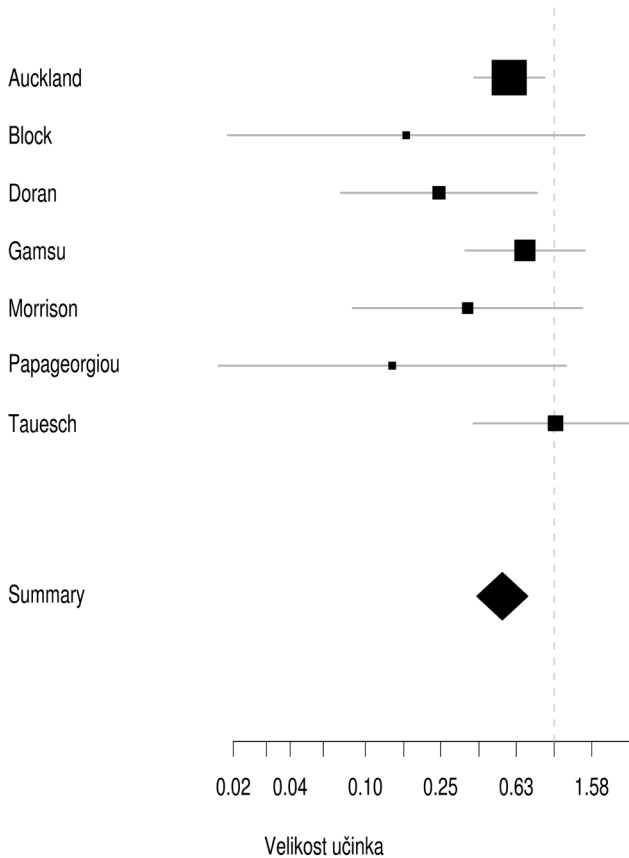
Vizualizacija podatkov se je z razmahom podatkovnega rudarjenja otesla priokusa nepotrebne okrasja in postala pomembna znanstvena disciplina. Vizualizacija podatkovnih struktur in rezultatov statističnih analiz je tako sestavni del ali pa celo končni cilj številnih statističnih orodij in metod. Metoda metaanalize se ponša z dvema vrstama grafik, po katerih je že ob bežnem prelistavanju znanstvene periodike tudi najbolj prepoznavna: drevesni diagram (angl. *forest plot*) je grafični prikaz vrednosti ocen velikosti učinkov posameznih študij ter njihove skupne ocene, lijakasti diagram (angl. *funnel plot*) pa dvorazsežni prikaz odnosa med velikostjo učinka študij in velikostjo njihovega raztrosa.

Drevesni diagram je ključni grafični povzetek rezultatov metaanalize (Slika 2). Idejo drevesnega diagrama je prvi predstavil Stephen Evans na osnovi razširjenega diagrama okvirja z ročaji (angl. *boxplot*; Lewis in Clarke, 2001). Velikost učinka posamezne študije predstavimo s kvadratom. Velikost kvadrata je odvisna od uteženosti študije, tj. od velikosti njenega vzorca. S horizontalno črto prikažemo interval zaupanja, v katerem z določeno stopnjo tveganja pričakujemo oceno. Običajno se uporablja 95-odstotni interval zaupanja. Skupno velikost učinka prikažemo z rombo. Interval zaupanja skupne ocene predstavljajo horizontalna oglišča romba. Ko imamo opravka z diskretnimi podatki in merami velikosti učinka, izpeljanimi iz verjetnosti dogodkov (npr. razmerja obetov), ocene velikosti učinka običajno logaritmiramo, s čimer zagotovimo simetričnost njihovih intervalov zaupanja (Agresti, 2002).

Lijakasti diagram je drugi standardni grafični prikaz v metaanalizi (Slika 3). Lijakasti diagram predstavlja odnos med velikostjo učinka študij in njihovimi ocenami raztrosa (Egger, Smith, Schneider in Minder, 1997). Uporablja se za odkrivanje sistematične heterogenosti (Song, Khan, Dinnes in Sutton, 2002). Pri popolni homogenosti je oblika diagrama simetrično lijakasta, asimetričnost pa nakazuje prisotnost heterogenosti, ki je lahko posledica različnih virov (omejenega izbora študij, vključenih v metaanalizo, slabe kakovosti študij, majhnih vzorcev, prave heterogenosti itd.).

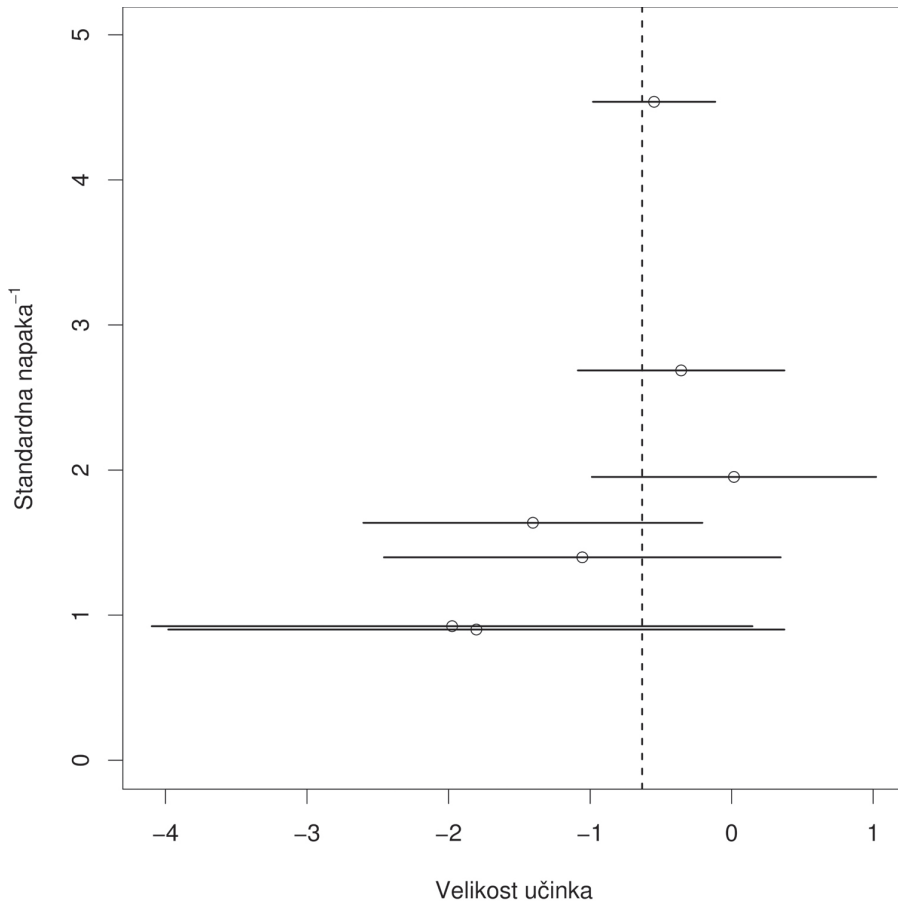
Kritika metaanalize

Po burnih začetnih odzivih Eysencka in raziskovalcev iz njegovega kroga je metaanaliza kot metoda tudi sama postala predmet znanstvenega proučevanja. Veliko podiplomskih študijskih programov, tako družboslovnih, humanističnih kot



Slika 2. Drevesni diagram. Diagram prikazuje velikosti učinkov za sedem različnih študij, ki so proučevale vpliv kortikosteroidne terapije na prezgodnji porod in neonatalno smrtnost. Surovi podatki so prosto dostopni na Cochranovem spletnem skladišču (<http://www.cochrane.org/>). Podatki so bili izvorno diskretni, mero velikosti učinka pa je predstavljalo razmerje obetov. Velikost učinka je prikazana v logaritemskih enotah. Velikost kvadratka je sorazmerna velikosti vzorca, na katerem je bila študija izvedena. Horizontalna črta označuje 95-odstotni interval zaupanja. Skupna ocena velikosti učinka je predstavljena z romбом, katerega levo in desno oglišče predstavljata 95-odstotni interval zaupanja skupne ocene. Skupna velikost učinka v našem primeru znaša 0,53, na osnovi česar lahko zaključimo, da je tveganje prezgodnjega poroda oz. neonatalne smrti pri nosečnicah s kortikosteroidnim zdravljenjem približno polkrat manjše kot pri kontrolni skupini. Bralec bo več informacij v zvezi z interpretacijo razmerja obetov našel v Agrestijevi monografiji (Agresti, 2002).

tudi naravoslovnih, v svojih predmetnikih že vključuje izbrana poglavja s področja metodologije metaanalize. Nekatera združenja organizirajo celo večdnevne študijske delavnice, na katerih študenti spoznajo uporabo metod metaanalize v različnih problematskih situacijah. Dandanes raziskovalci z različnih področij proučujejo uporabo



Slika 3. Lijakasti diagram. Diagram prikazuje velikost učinka študij v odnosu do mer njihovega raztrosa. Horizontalna črta označuje 95-odstotni interval zaupanja posameznih študij. Oblika diagrama je približno simetrična, kar nakazuje, da je homogenost študij zadovoljiva.

različnih mer velikosti učinka nad različnimi tipi podatkov, proučujejo veljavnost mer ocenjevanja heterogenosti med študijami in razvijajo nove računske metode za povzemanje skupnega učinka (npr. Bayesovi modeli). Na splošno lahko kritike metode metaanalize razčlenimo v treh sklopih.

Metaanaliza zahteva, da so podatki, s katerimi vstopamo v analizo, med seboj primerljivi. V metodoloških učbenikih je ta problem opisan z metaforo združevanja jabolk in pomaranč. Nehomogenost vzorcev, različni tipi eksperimentalnih zasnov in vsebinsko močno različna operacionalizacija merskih spremenljivk v posameznih študijah so glavni viri sistematičnih napak lahkovernih (meta)raziskovalcev.

Drug glavni problem je kvaliteta študij, vključenih v metaanalizo. Problem izbora kvalitetnega in reprezentativnega vzorca je ključnega pomena ne samo za

ustrezno statistično oceno proučevanega fenomena, pač pa tudi za generalizabilnost zaključkov. Stabilnost skupnega učinka, izračunanega v metaanalizi, je smiselno preveriti s pomočjo analize občutljivosti. Pri analizi občutljivosti gre za to, da opazujemo raztros skupne ocene velikosti učinka v odvisnosti od nabora študij, vključenih v metaanalizo, od vrste uporabljene metode za izračun skupne ocene, eksperimentalne zasnove itd. Pri dobro zasnovanih metaanalizah je tako pričakovati, da se z izločitvijo študij z majhnim številom udeležencev skupna ocena ne bo statistično značilno spremenila.

Uredniška politika znanstvenih časopisov je v veliki večini primerov taka, da raziskovalci praviloma lažje objavijo raziskavo, ki poroča o statistično značilnih rezultatih, kot pa raziskavo, ki do takih izsledkov ni prišla. Kljub temu, da obstajajo namenske revije, ki objavljajo prav take študije (npr. *Journal of Negative Results in BioMedicine*), jih večina pristane na smetišču zgodovine. Heterogenosti, ki iz tega izhaja, pravimo pristranost v objavljanju (angl. *publication bias*). Zato je ključnega pomena, da v metaanalizo poleg klasičnih znanstvenih člankov, ki jih indeksirajo mednarodne bibliografske zbirke (npr. PsycINFO, Medline), vključimo tudi pregledne članke, povzetke in programe konferenc in simpozijev, diplomske, magistrske in doktorske naloge, raziskovalna poročila in elaborate, neindeksirane strokovne članke, poročila vladnih in nevladnih organizacij in bibliografije (Rothstein, Sutton in Borenstein, 2005).

Programska oprema

Pestrost numeričnih postopkov za analizo zbranih podatkov na različnih tipih merskih lestvic narekuje razvoj specializirane programske opreme, ki tudi manj večjemu uporabniku omogoča kvalitetno in metodološko ustrezno izvedbo metaanalize. Statistično bolj podkovani raziskovalci lahko metaanalizo opravijo z uporabo splošnih statističnih programskih paketov (npr. SPSS, SAS) ali celo z uporabo elektronske preglednice (npr. Excel, OpenOffice Calc), medtem ko večina raziskovalcev raje poseže po namenski programski opremi. Izbor nekaterih programov je povzet v Tabeli 1. Med njimi najdemo tako komercialne kot tudi prostodostopne in odprtokodne programe. Glavna slabost vseh programov je razmeroma dolg čas ogrevanja, saj za osvojitve osnovnih principov manipulacije s podatki običajno potrebujemo dalj časa kot za samo izvedbo metaanalize. Z namenom enostavne, hitre in učinkovite podpore k metaanalitični metodologiji smo zato razvili spletni strežnik RMetaWeb, ki uporabniku ponuja možnost interaktivne online analize podatkov. Gre za prvo tovrstno spletno orodje in predstavlja močno alternativo obstoječi programski opremi. Numerično jedro spletnega strežnika predstavlja okolje R za statistično analizo in grafiko (R Development Core Team, 2008).

R sta pred dobrim desetletjem zasnovala Ross Ihaka in Robert Gentleman z Univerze v Aucklandu, Nova Zelandija (Ihaka in Gentleman, 1996). V desetih letih je R postal vodilni statistični programski paket in programski jezik. Gre za

Tabela 1. Programska podpora za meta analizo.

Program	URL naslov
Comprehensive Meta-Analysis*	http://www.meta-analysis.com/
EasyMA	http://www.spc.univ-lyon1.fr/easyma.dos/
EpiMeta	http://ftp.cdc.gov/pub/Software/epimeta/
Hepima	http://www.hsph.harvard.edu/faculty/spiegelman/tcs.html
Meta-Analysis 5.3	http://userpage.fu-berlin.de/~health/meta_e.htm
Meta-Analyst	http://www.medepi.net/meta/MetaAnalyst.html
Meta-Stat	http://edres.org/meta/metastat.htm
MetaWin*	http://www.metawinsoft.com/
MIX	http://www.mix-for-meta-analysis.info/
RevMan	http://www.cc-ims.net/RevMan
WEasyMA*	http://www.weasya.com/
RMetaWeb	http://www2.arnes.si/~akastr1/

Opomba: Z zvezdico (*) so označeni plačljivi programi.

odprtokodno (in torej brezplačno) implementacijo predmetnega jezika S, ki so ga zasnovali v Bellovih laboratorijih. Bistvo okolja R predstavlja njegova nadgradljivost s programskimi paketi. Osnovna namestitvena distribucija vključuje le osnovni nabor paketov, ostale pakete pa si uporabnik namešča po potrebi preko omrežja zrcalnih strežnikov CRAN (angl. *The Comprehensive R Archive Network*). Okolje R je na voljo za različne izvedbe operacijskih sistemov (Unix, Linux, Mac OS, Windows) in ponuja zelo obsežno programsko podporo za delo s podatki, računske operacije in grafične prikaze. Njegove glavne prednosti v primerjavi z ostalimi, bolj razširjenimi in poznanimi statističnimi programi so predvsem: (i) visoka razvitost, razumljivost in preglednost programskega jezika, ki vključuje široko paleto podatkovnih struktur, zanke, rekurzivne klice in pogojno izvajanje programske kode; (ii) hitre računske operacije nad vektorji in matrikami; (iii) učinkovito shranjevanje in priklic podatkov; (iv) zmogljivi grafični podsistemi in paketi za vizualizacijo podatkov ter (v) neomejena razširljivost, ki omogoča integracijo z drugimi programskimi jeziki (Fortran, C, C++, Java, Perl, Python).

Poleg enostavnega dostopa do klasičnih orodij za analizo in vizualizacijo eno- in dvodimenzionalnih podatkov ponuja R zahtevnejšemu raziskovalcu s področja psihologije tudi celo paleto orodij za napredne statistične analize. Pred kratkim so bili razviti paketi, ki omogočajo preprost dostop do funkcij za analizo modelov latentnih potez, strukturnega modeliranja in korespondenčne analize. Metaanalizi sta namenjena dva, med seboj razmeroma podobna paketa: *rmeta* in *meta*. Funkcije slednjih dveh smo uporabili tudi za razvoj spletnega strežnika RMetaWeb (<http://www2.arnes.si/~akastr1/>).

Spletni strežnik RMetaWeb je v celoti razvit z uporabo prostih in odprtokodnih orodij. Dinamičen pretok med numeričnim delom in spletnim prikazom je implementiran s pomočjo paketa CGIwithR (Firth, 2003). Uporabnik datoteko s surovimi podatki najprej pretoči na spletni strežnik in označi strukturo vhodne datoteke. Trenutno podprti formati vhodnih podatkov so standardne tekstovne datoteke (tsv, csv), excelove (xls) ali SPSS (sav) datoteke. V drugem koraku izbere želeno mero velikosti učinka ter metodo združevanja rezultatov. Strežnik nato izpiše rezultate metaanalize za dane podatke ter ponudi možnost prenosa rezultatov v obliki tekstovne datoteke, ki si jo uporabnik lahko shrani za kasnejšo uporabo. V tretjem koraku sistem izriše drevesni in lijakasti diagram. Odvisno od namena uporabe lahko velikost slike poljubno spreminjamo. Uporabniku je na voljo tudi možnost prenosa diagramov v PNG in PDF zapisu. PNG je enoznačen slikovni format in je namenjen predvsem ogledu slik na spletu ali predstavitvam, medtem ko je slika v PDF formatu shranjena v vektorski obliki in jo lahko uporabimo neposredno za tisk.

V trenutni implementaciji RMetaWeb omogoča analizo zveznih in diskretnih spremenljivk preko različnih mer velikosti učinka za statistične modele s stalno in slučajno eksperimentalno zasnovo. Odvisno od izkušenosti in potreb, uporabnik kombinira različne mere velikosti učinka in metode za izračun njene skupne ocene. Pri diskretnih podatkih lahko uporabnik izbira med merami razmerja obetov, razlik v tveganju (oz. ogroženosti) in relativnim tveganjem, pri zveznih podatkih pa je trenutno na voljo Hedgesova g statistika. Za izračun skupne velikosti učinka so na voljo tri standardne metode: Mantel-Haenszelova, Petova in metoda inverzne variance.

V nadaljevanju bomo implementirali še algoritem za analizo občutljivosti posameznih študij ter ponudili interaktivne obrazce za pretvarjanje različnih mer velikosti učinka in izračunavanje njihovih intervalov zaupanja.

Zaključek

Newton je leta 1676 v pismu Hooku zapisal: "Če sem videl dlje, je to zaradi tega, ker sem stal na ramenih velikanov." Z uvedbo metaanalize (ali pa njenim ponovnim odkritjem) je psihologija za skoraj dvajset let prehitela razvoj danes zelo opevanega področja odkrivanja zakonitosti iz podatkov. Upali bi si celo trditi, da je metoda metaanalize eden od temeljni kamnov vseh sodobnih tehnologij znanja.

Metaanaliza sicer zahteva razmeroma velik vložek dela in premišljeno kombiniranje kvalitativne in kvantitativne analize, vendar po drugi strani omogoča sprotno in sistematično spremljanje najnovejših znanstvenih spoznanj, učinkovitejšo izrabo obstoječih podatkov, pomembno prispeva h kakovosti obstoječega znanja o določenem proučevanem fenomenu in nenazadnje nudi podporo pri gradnji novih raziskovalnih domnev.

Uporaba katerekoli statistične metode zahteva veliko znanja in izkušenj. Zavedati se moramo, da metaanaliza ni nadomestek raziskovalčeve ustvarjalnosti, pač pa le orodje, ki lahko močno pospeši in izboljša kvaliteto raziskovalnega dela.

Literatura

- Agresti, A. (2002). *Categorical data analysis* (2. izd.). Hoboken, NJ: Wiley.
- Balding, D. J. (2006). A tutorial on statistical methods for population association studies. *Nature Reviews Genetics*, 7(10), 781–791.
- Cankar, G. in Bajec, B. (2003). Velikost učinka kot dopolnilo testiranju statistične pomembnosti razlik [Effect size as a supplement to statistical significance testing]. *Psihološka obzorja*, 12(2), 97–112.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. izd.). Hillsdale, NJ: Lawrence Erlbaum.
- Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, 49(12), 997–1003.
- DerSimonian, R. in Kacker, R. (2007). Random-effect model for meta-analysis of clinical trials: An update. *Contemporary Clinical Trials*, 28(2), 105–114.
- DerSimonian, R. in Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7(3), 177–188.
- Egger, M., Ebrahim, S. in Smith, G. D. (2002). Where now for meta-analysis? *International Journal of Epidemiology*, 31(1), 1–5.
- Egger, M., Smith, G. D., Schneider, M. in Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, 315(7109), 629.
- Eysenck, H. J. (1952). The effect of psychotherapy: An evaluation. *Journal of Consulting Psychology*, 16(5), 319–324.
- Eysenck, H. J. (1965). The effects of psychotherapy. *International Journal of Psychiatry*, 1, 97–142.
- Eysenck, H. J. (1978). An exercise in mega-silliness. *American Psychologist*, 33(5), 517.
- Ferligoj, A., Leskošek, K. in Kogovšek, T. (1995). *Zanesljivost in veljavnost merjenja [Reliability and validity of measurement]*. Ljubljana: Fakulteta za družbene vede.
- Field, A. P. (2003). The problems in using fixed-effect models of meta-analysis on real-world data. *Understanding Statistics*, 2(2), 105–124.
- Firth, D. (2003). CGIwithR: Facilities for processing web forms using R. *Journal of Statistical Software*, 8(10), 1–8.
- Fisher, R. A. (1970). *Statistical methods for research workers*. New York: MacMillan.
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5(10), 3–8.
- Gordon, M. H., Loveland, E. H. in Cureton, E. E. (1952). An extended table of chi-square for two degrees of freedom, for use in combining probabilities from independent samples. *Psychometrika*, 17(3), 311–316.
- Hedges, L. V. in Olkin, I. (1985). *Statistical methods for meta-analysis*. Boston, MA: Academic Press.

- Hill, G. B. (2000). Archie Cochrane and his legacy. An internal challenge to physicians' autonomy? *Journal of Clinical Epidemiology*, 53(12), 1189–1192.
- Hunter, J. E. in Schmidt, F. L. (2000). Fixed effects vs. random effects meta-analysis: Implications for cumulative research knowledge. *International Journal of Selection and Assessment*, 8(4), 275–292.
- Ihaka, R. in Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5(3), 299–314.
- Kisamore, J. L. in Brannick, M. T. (2008). An illustration of the consequences of meta-analysis model choice. *Organizational Research Methods*, 11(1), 35–53.
- Lewis, S. in Clarke, M. (2001). Forest plots: Trying to see the wood and the trees. *British Medical Journal*, 322(7300), 1478–1480.
- Mantel, N. in Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 22(4), 719–748.
- McNemar, Q. in Terman, L. M. (1936). Sex differences in variational tendency. *Genetic Psychology Monographs*, 18(1), 31.
- Olkin, I. (1995). Statistical and theoretical considerations in meta-analysis. *Journal of Clinical Epidemiology*, 48(1), 133–146.
- Pearson, K. (1904). Report on certain enteric fever inoculation statistics. *British Medical Journal*, 2(2288), 1243–1246.
- Plackett, R. L. (1958). Studies in the history of probability and statistics: VII. The principle of the arithmetic mean. *Biometrika*, 45(1–2), 130–135.
- R Development Core Team (2008). *R: A language and environment for statistical computing* [Računalniški software]. Dunaj, Avstrija. (ISBN 3-900051-07-0)
- Rosenthal, R. (1976). *Experimenter effect in behavioral research* (razširjena izd.). New York: Halsted Press.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Newbury Park, CA: Sage.
- Rothstein, H. F., Sutton, A. J. in Borenstein, M. (2005). *Publication bias in meta-analysis: Prevention, assessment and adjustment*. London: Wiley.
- Sackett, D. L., Rosenberg, W. M., Gray, J. A., Haynes, R. B. in Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn't. *British Medical Journal*, 312(7023), 71–72.
- Schmidt, F. L. (1992). What do data really mean? Research findings, meta-analysis, and cumulative knowledge in psychology. *American Psychologist*, 47(10), 1173–1181.
- Sedlmeier, P. in Gigerenzer, G. (1989). Do studies of statistical power have an effect on the power of studies? *Psychological Bulletin*, 105(2), 309–316.
- Smith, M. L. in Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 32(9), 752–760.
- Song, F., Khan, K. S., Dinnes, J. in Sutton, A. J. (2002). Asymmetric funnel plots and publication bias in meta-analysis of diagnostic accuracy. *International Journal of Epidemiology*, 31(1), 88–95.
- Tippet, L. H. C. (1931). *The methods of statistics*. London: Williams & Norgate.
- Torgerson, C. (2003). *Systematic reviews*. London: Continuum.

- Wachter, K. W. (1988). Disturbed by meta-analysis. *Science*, 241(4872), 1407–1408.
- Whitehead, A. (2002). *Meta-analysis of controlled clinical trials*. West Sussex: Wiley.
- Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis*. Newbury Park, CA: Sage.
- Wood, J. (2008). Methodology for dealing with duplicate study effects in a meta-analysis. *Organizational Research Methods*, 11(1), 79–95.

Gender differences in child/adolescent personality traits: Slovenes and Russians compared[#]

Maja Zupančič^{1}, Helena R. Slobodskaya² and Gennady G. Knyazev²*

¹*Department of Psychology, University of Ljubljana, Ljubljana, Slovenia*

²*State Research Institute of Physiology, Siberian Branch of the Russian Academy of Medical Sciences,
Novosibirsk, Russia*

Abstract: The present study compared gender differences in parental perceptions of child personality trait expression across different age groups in Slovenia and Russia. The patterns of gender differences in adolescents were further explored using parent and self-rating methods of assessment. Employing the Inventory of Child Individual Differences (Halverson et al., 2003) reports on large samples of children/adolescents (2 to 15 years) in both countries were obtained. Gender differences were small and similar across countries, and their patterns were closer to each other for the same data source across countries than they were for the two sources of information within the country. The parents in both countries concurred in rating daughters as somewhat more achievement orientated, compliant, considerate, and organized, whereas sons were assessed as slightly more active, antagonistic, and distractible. At the higher-order trait level, girls were perceived to be more conscientious and agreeable relative to the boys, especially from middle childhood onwards. Regarding self-reports, adolescent girls also scored significantly higher on Considerate and Positive Emotions scales than boys, while the former reported on higher agreeableness relative to the latter in Slovenia only. Gender differences were evident even in parental assessments of the youngest group but they did not systematically increase over children's age.

Keywords: personality, children, adolescents, gender differences, cross cultural differences

Razlike med spoloma v osebnostnih potezah otrok/mladostnikov: primerjava Slovencev in Rusov

Maja Zupančič¹, Helena R. Slobodskaya² in Gennady G. Knyazev²

¹*Oddelek za psihologijo, Univerza v Ljubljani, Ljubljana*

²*State Research Institute of Physiology, Siberian Branch of the Russian Academy of Medical Sciences,
Novosibirsk, Russia*

Povzetek: V raziskavi smo primerjali razlike med spoloma, kot se odražajo skozi starševo zaznavo izraznosti osebnostnih potez pri različno starih skupinah slovenskih in ruskih otrok/mladostnikov. Vzorce razlik med spoloma v zgodnjem mladostništvu smo preučevali tudi z dvema metodama, s starševo oceno

* Naslov/Address: Maja Zupančič, Faculty of Arts, Department of Psychology, Aškerčeva 2, 1000 Ljubljana, Slovenia, e-mail: maja.zupancic@ff.uni-lj.si

[#] Acknowledgements: This study was supported by grants of the Russian Foundation for Basic Research (RFBR) № 08-06-00011-a and № 08-06-00016-a and Russian Humanitarian Foundation № 07-06-00016-a. The support in Slovenia came from the Slovenian Research Agency № J5-6450-0581-04 and № P5-0062, and the Slovenian Ministry for Education and Sports № V5-0544-01 and № V5-0935-04.

osebnosti njihovih otrok in s samooceno mladostnikov. Z Vprašalnikom medosebnih razlik med otroki (Halverson idr., 2003) smo zbrali podatke o velikih vzorcih ciljnih posameznikov (starih od 2 do 15 let) v Sloveniji in Rusiji. Razlike med spoloma so bile majhne in podobne v obeh deželah. Vzorci teh razlik, ugotovljeni z isto metodo pridobivanja podatkov, so bili bolj podobni med deželama kot vzorci razlik, ki so v posamezni deželi temeljili na podatkih, zbranih z različnima metodama. Slovenski in ruski starši so na splošno hčerkam pripisali nekoliko višjo raven usmerjenosti k dosežku, ugodljivosti, obzirnosti in organiziranosti kot fantom, ki so jih ocenili kot nekoliko bolj dejavne, nasprotnovalne in odkrenljive. Na ravni robustnih osebnostnih potez pa so dekleta opisali kot bolj vestna in sprejemljiva v primerjavi s fanti, zlasti od obdobja srednjega otroštva dalje. Mladostnice so se v obeh deželah ocenile značilno višje glede obzirnosti in pozitivnega čustvovanja v primerjavi z mladostniki, medtem ko so prve le v Sloveniji poročale o značilno višji ravni sprejemljivosti kot fanti. Razlike med spoloma so se pokazale že v starševih ocenah najmlajše skupine otrok, vendar se niso sistematično povečevale s starostjo ciljnih posameznikov.

Ključne besede: osebnost, otroci, mladostniki, razlike med spoloma, medkulturne razlike

CC = 2840

Personality traits are often considered as consistent and enduring tendencies to think and act in characteristic ways that are more expressions of human biology than products of life experience (McCrae & Costa, 2003; McCrae et al., 2000). The Five-Factor Model (FFM), a widely accepted taxonomy of personality traits, is one of universal generalizations, which at the time is enormously influential in the personality research. The FFM classifies most traits in terms of five broad dispositions: Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. The FFM has been supported in a large variety of cultures using translations of personality questionnaires in both self-report (McCrae & Costa, 2003) and observer rating (McCrae, Martin, & Costa, 2005) versions. Studies of personality within the FFM framework also repeatedly noted differences in personality trait expression across cultures and gender (e.g., Costa, Terraciano, & McCrae, 2001; McCrae, Costa, Martin et al., 2004; McCrae et al., 2000; McCrae et al., 2002; McCrae, Terraciano et al., 2005).

Measurement of personality traits in childhood and adolescence

It has been argued that the basic structure of personality is basically the same in different adolescent and adult age groups (Allik, Laidra, Realo, & Pullman, 2004; McCrae et al., 2002; McCrae, Terracciano et al., 2005). On the other hand, developmental psychologists typically refer to stable individual differences among children as evidence of temperament. The traditions of personality and temperament research differ substantially (McCrae et al., 2002) and a consensus regarding the structure of children's personality starts to emerge within the framework of the FFM (McCrae et

al., 2000; Shiner & Caspi, 2003). The Five-Factor structure of personality has been recovered from ratings by parents and teachers of children and early adolescents (e.g., Goldberg, 2001; Halverson et al., 2003; Knyazev, Zupančič, & Slobodskaya, 2008; Mervielde, Buyst, & DeFruyt, 1995; Mervielde & DeFruyt, 1999; Resing, Bleichrodt, & Dekker, 1999; Zupančič, Gril, & Kavčič, 2006).

Adolescents' self-ratings have also revealed a factor structure similar to adults' self-ratings of personality (Allik et al., 2004; Donahue, 1994; McCrae et al., 2002; Scholte, van Aken, & van Lieshout, 1997; van Lieshout & Haselager, 1994). Furthermore, Markey, Markey, Ericksen and Tinsley (2002) have found that pre-adolescent children are able to reliably rate themselves using a standard measure of adult personality (NEO–Five-Factor Inventory; NEO-FFI) and that parent ratings of their children's personalities are related to these self-reports. Similar findings were shown with samples of older (Barbaranelli, Caprara, Rabasca, & Pastorelli, 2003) and younger school-age children (Measelle, John, Ablow, Cowan, & Cowan, 2005), and early adolescents (Knyazev et al., 2008), using measures specifically designed for child self-reports. Moreover, conducting confirmatory factor analyses with data obtained by one of such assessment tools, i.e. the Inventory of Child Individual Differences (ICID, Halverson et al., 2003), Knyazev et al. (2008) demonstrated that the structure of child/adolescent personality is similar in Slovenia and Russia, across genders, ages, and informants (parents and adolescents). The invariance of the proposed model, i.e. Extraversion (sociable, positive emotions, activity), Conscientiousness (achievement, organized, un-distractible), Neuroticism (fearful/insecure, shy, negative affect), Disagreeableness (antagonism, strong willed), and Openness (intelligent, open to experience), therefore allows to conduct between group analyses (e.g., gender differences) in the perceived higher- and lower-order trait expression (the five factors) of target children across age, country and methods of assessment (e.g., parent vs. self-report).

Gender differences in adult and child personality

Maccoby and Jacklin (1974) conducted the first major review of research on sex-related differences in cognition, temperament/personality, and social behavior in children and adults. They showed that males are consistently perceived as more assertive and less anxious than females, but that gender differences are small. Recent work suggested similar conclusions. In general, gender differences in traits related to Neuroticism have been reported, with females scoring higher than males (e.g., Feingold, 1994; Lynn & Martin, 1997; McCrae, Terracciano et al., 2005). In comparison to females, males have been also found to score higher on Psychoticism (Lynn & Martin, 1997), assertiveness, excitement seeking (Extraversion facets), ideas (Openness facet), and lower on Agreeableness (especially Warmth), esthetics and feelings (Openness facets) (McCrae, Terracciano et al., 2005). These results were invariant

across countries, methods (self- and report by others) and adult age-groups (Costa et al., 2001; Feingold, 1994; Lynn & Martin, 1997; McCrae, Terracciano et al., 2005).

Studies on gender differences in personality over childhood are scarce as child personality was long considered in terms of temperament. A recent meta-analysis (Else-Quest, Hyde, Goldsmith, & Van Hulle, 2006) of gender differences in temperament (3-month- to 13-year-olds) suggests a large gender difference in effortful control and moderate differences in its lower order traits (attention regulation, inhibitory control, perceptual sensitivity), all favoring girls. Surgency and some traits (e.g., activity, impulsivity, high-intensity pleasure) within this factor showed a small difference favoring boys, while approach, positive mood, and shyness showed very small effect sizes favoring girls. With regard to negative affectivity, very small effects were found for difficult and intensity favoring boys and for fear favoring girls.

Similarly to the temperament research findings on gender differences in effortful control (Else-Quest et al., 2006) the most robust gender effect for Big Five categories in the analysis of free parental descriptions of 3 to 12 year old children from seven countries was that girls were ascribed more Conscientiousness descriptive words than boys (De Fruyt, Van Hiel, & Buyst, 1998). Contrary to the adult studies using self- or observer reports (Costa et al., 2001; Feingold, 1994; Lynn & Martin, 1997; McCrae, Terracciano et al., 2005), but in line with free descriptive approach, remarkable similarities across countries were also revealed for ICID ratings of child/early adolescent Conscientiousness with girls scoring higher than boys. At the lower order trait level, girls were consistently ascribed more organization, achievement orientation and compliance than boys who, in turn, scored higher on distractibility and activity than girls (e.g., Halverson, 2003; Slobodskaya, 2005; Zupančič et al., 2006). However, several gender differences were country specific. Slovene parents, for example, perceived girls as more emotionally stable than boys (Zupančič et al., 2006), especially due to lower level of fear/insecurity of the former in comparison to the latter, a result consonant with findings of a longitudinal, multiple-informant ICID study conducted with parents of Slovene pre-schoolers and first-graders (Zupančič & Kavčič, 2005, 2007).

In general, results obtained by cross-sectional studies of parent inventory reports (ICID) on children and early adolescents in different countries (Halverson, 2003; Slobodskaya, 2005; Zupančič et al., 2006) suggest small gender differences across different countries, especially in Slovenia and Russia. Previous findings also indicate that gender differences in mean levels of infant temperament traits and adolescent personality traits in Russia were smaller by comparison with Western society (Gartstein, Slobodskaya, & Kinsht, 2003; Slobodskaya, Safronova, Knyazev, & Wilson, 2001).

Three general models to address proximal causes of gender differences were suggested (see e.g. Feingold, 1994). The *biological model* proposes that the observed gender differences in personality scores reflect innate temperamental gender differences and emphasizes biological basis underlying individual differences in personal-

ity. According to the *socio-cultural model*, social and cultural factors directly produce gender differences in personality. There are several examples of this model: (a) the social role model contends that gender differences in social behavior stem from gender roles which determine behaviors that are appropriate for each gender and these behaviors may shape personality; (b) the expectancy model posits that socio-cultural factors eventuate in gender stereotypes. These cause gender differences in personality because the stereotype holders treat others in ways that result in others conforming to these stereotypes; (c) the artifact model explains gender differences on personality scale scores rather than in underlying constructs. Socio-cultural factors, e.g. gender stereotyping, result in genders holding different perspectives about the importance of possessing various personality traits and these differences differentially bias self-(other-)reports on personality characteristics. The gender differences in personality scale scores thus do not reflect solely corresponding gender differences in personality constructs that the inventory is supposed to measure. Finally, the *biosocial model* proposes that biological and socio-cultural factors both cause gender differences in personality. If genders are perceived differently because of the observable behavioral differences that are linked to innate temperamental gender differences, genders may be treated differently. And if social and cultural factors affect personality development, these factors may augment inherent gender differences.

The present study

This study aimed at comparison of gender differences in mean levels of higher- and lower-order child/adolescent personality traits in Slovenia and Russia, the countries differing in size, political-geographic position, exposure to Western influences, and economical situation but similar in some aspects of recent history and in scores of dimensions of culture as proposed by Hofstede (2001). Since most of previous child studies relied exclusively on parent reports, some of the gender differences in the perceived trait expression might have emerged due to different perceptions of their sons and daughters by parents in different cultures. One way to partially overcome these response biases would be a comparison of patterns of cross-cultural gender differences in parent- versus self-reported data. Therefore, we collected cross-sectional data from two sources: adult reports on children¹ and adolescent self-ratings. To reveal the robust gender differences across the two countries, we compared parental ratings in children of different ages (2 to 15 years). According to the biosocial hypothesis, the gender differences would appear even in toddlers and they would tend to increase as the children develop.

Although most of the cultures show the same pattern of relatively small gender differences, there are cultural variations in the magnitude of these differences. Con-

¹ The term children is used for parent ratings to designate children as well as early adolescents.

trary to the prediction from the social role model, gender differences in adults are the most pronounced in relatively rich, individualist, egalitarian and masculine cultures (Costa et al., 2001; McCrae, Terracciano et al., 2005). With respect to the dimensions of culture (Hofstede, 2001), both Slovenia and Russia appear to be relatively collectivist, high on power distance and low on masculinity. Accordingly, the personality gender differences in adult samples of the two countries as compared to those in most Western societies were found smaller and not consistently significant across the five factors (McCrae, Terracciano et al., 2005). Considering this along with previous reports on personality gender differences in children (Halverson, 2003; Slobodskaya, 2005; Zupančič et al., 2006) and given that Slovenia is, relative to Russia, higher on Gross National Income (World Bank, 2004) but lower on power distance, individualism, and masculinity (Hofstede, 2001), we expected to find similar and small gender differences in the two countries. We hypothesized that girls would be perceived to express higher levels of Agreeableness and Conscientiousness, and the corresponding marker traits (particularly according to parental ratings) than boys. Drawing from consistent findings in adults (Costa et al., 2001; Feingold, 1994; McCrae, Terracciano et al., 2005), it was also assumed that gender differences in Neuroticism in favor of females would start to emerge as children move toward adulthood.

Method

Slovene samples

The samples were selected from different Slovene regions covering urban and rural areas. The children attended one of the public (pre)schools and came from diverse family educational background.

Sample 1. Altogether 1850 adults (mostly mothers) rated 893 boys and 957 girls aged 2 to 15 years ($M = 8.7$, $SD = 4.1$); 10.5% of the mothers and 9% of the fathers completed compulsory school, 21.7% of the mothers and 34.8% of the fathers finished vocational school (10 or 11 years of schooling), 27.9% of the mothers and 25.9% of the fathers completed high school (12 or 13 years), 18.5% of the mothers and 13% of the fathers had additional two or three years of higher education, 16.7% of the mothers and 11.7% of the fathers had university degree, and 4.5% of the mothers and 5.4% of the fathers had academic degree (MA or PhD). The sample was divided into four age groups (data on age was missing for 17 children): 2–3 years ($N = 397$; 49% boys and 51% girls), 4–6 years ($N = 300$; 50% boys and 50% girls); 7–11 years ($N = 620$; 48% boys and 52% girls) and 12–15 years ($N = 516$; 47% boys and 53% girls).

Sample 2. A sample of 393 youths (186 boys, 207 girls) aged 12 to 14 years ($M = 13.5$, $SD = 0.33$) completed the ICID; 12.4% and 10.6% of adolescents' mothers and fathers, respectively, completed compulsory school, 22.1% and 34.3% of the mothers and fathers, respectively, finished vocational schooling, 29.3% and 27.9% of

the mothers and fathers, respectively, graduated from high school, 16.6% and 8.2% of the mothers and fathers, respectively, had higher education, 14.3% and 12.8% of the mothers and fathers, respectively, graduated from university, while 5.4% and 6% of the mothers and fathers, respectively, obtained academic degree.

Russian samples

The data were collected both in urban and rural areas of western Siberia, Russia. Most data came from Novosibirsk, the third largest Russian city and the economic and academic capital of Siberia. The recruitment was aimed to collect data from the diverse socioeconomic background.

Sample 1. Altogether 1635 parents or caregivers rated 844 boys and 791 girls aged 2 to 15 years ($M = 9.5$, $SD = 3.8$); 1.7% of the mothers and 1.7% of the fathers had eight years or less of compulsory schooling, 8.8% of the mothers and 14.8% of the fathers completed ten years of schooling, 42.2% of the mothers and 39.1% of the fathers had additional two or three years of college education, 45.9% of the mothers and 40.6% of the fathers had five or six years at the university or institute and 1.5% of the mothers and 3.8% of the fathers had an academic degree. For the subsequent analysis the sample was divided into four age groups (data on age was not reported for 29 children): 2–3 years ($N = 104$; 54% boys and 46% girls), 4–6 years ($N = 320$; 57% boys and 43% girls), 7–11 years ($N = 660$; 54% boys and 46% girls), and 12–15 years ($N = 552$; 52% boys and 48% girls).

Sample 2. A sample of 553 Russian youths (259 boys, 294 girls) aged 12 to 14 years ($M = 13.1$, $SD = 0.81$) completed questionnaires. Unfortunately, we do not have data on parental education for adolescents' reports, but according to parents' ratings their education correlated with 4 out of 15 traits and all of these correlations were low (data available at request).

Instrument

The Inventory of Child Individual Differences (ICID; Halverson et al., 2003) is a comprehensive, age and culture neutral measure of child personality consisting of 108 adjectives and phrases that describe children's characteristics in parental natural language. Each item is rated using a seven-point Likert scale with responses ranging from "much less than the average child or not at all" to "much more than in the average child". The items form 15 mid-level (lower-order) scales capturing aspects of the child's cognitive, social and emotional behavioral tendencies. The scales measure the following constructs: achievement orientation, activity level, antagonism, compliance, consideration, distractibility, fear/insecurity, intelligence, negative affectivity, openness to experience, organized behavior, positive emotionality, shyness, sociability, and strong will. For parent ratings, the scales have shown good reliabilities across countries and age groups (Halverson et al., 2003; Knyazev

& Slobodskaya, 2005; Knyazev et al., 2008; Zupančič et al., 2006). A close correspondence of the ICID scales to the set of FFM markers reported by other authors (e.g., Goldberg, 2001; Mervielde & De Fruyt, 1999) was demonstrated and so were the scales' links to measures of temperament and problem behavior (Halverson et al., 2003).

Both the Russian and the Slovene version of the ICID lower- and higher-order scales (the superfactors, Extraversion, Conscientiousness, (dis)Agreeableness, Neuroticism, and Openness) were extensively investigated. Sound psychometric properties, such as internal reliability, cross-observer agreement, stability over time, concurrent and longitudinal predictive validity against measures of social functioning, family relationships, academic achievement, and problem behaviour were suggested (Kavčič & Zupančič, 2006; Knyazev & Slobodskaya, 2005; Slobodskaya, 2005; Zupančič & Kavčič, 2007; Zupančič, Sočan, & Kavčič, 2007).

Procedure

The informants of the four samples participated on a voluntary basis and they were fully informed on the objectives and methods of the study. All of the participants were assured in the confidentiality of their responses. In Russia, the study was approved by the State Research Institute of Physiology SB RAMS Ethical Committee, while in Slovenia the required parental consent was obtained for both samples.

Across the countries, both the convenience and randomized sampling of parents and adolescents were used. The parents received envelopes containing the ICID with short written instructions on how to fill-in an inventory (demographic data was also requested). The parents were asked to complete the ICID in their home and return it in sealed envelopes within two weeks.

In both countries, most of adolescents were recruited through schools, self-reports were administered in school time as a part of the regular curriculum. The inventories which also contained brief questions on demographics were completed in approximately half an hour.

Results

Tables A1 (the Slovene sample) and A2 (the Russian sample) in the Appendix display means and standard deviations in ICID lower- and higher-order scale scores in boys and girls across four age groups (toddlerhood, early childhood, middle and late childhood, and early adolescence). In addition, Table 1 presents means and standard deviations in adolescents' ICID self-ratings for lower- and higher-order traits across genders and across the two countries. The scoring-key for the higher-order traits (the Big Five) utilised in the present study relied on the model that fitted data well in both countries, across age groups, genders, and rating methods (parent vs. self-

Table 1. Means (and standard deviations in parentheses) in ICID adolescent self-ratings: Boys and girls across the two countries.

ICID Scale	Slovenia (N = 393)		Russia (N = 553)	
	Boys	Girls	Boys	Girls
Achievement Orientation	5.01 (0.92)	4.94 (0.92)	4.69 (0.97)	4.63 (0.97)
Activity Level	5.05 (1.05)	4.85 (1.03)	4.89 (1.07)	4.68 (1.07)
Antagonism	3.02 (0.84)	2.74 (0.78)	2.78 (1.02)	2.60 (1.04)
Compliant	4.67 (0.88)	4.67 (0.77)	4.54 (1.00)	4.51 (0.81)
Considerate	4.54 (0.89)	4.88 (0.80)	4.71 (1.01)	4.92 (0.92)
Distractible	3.49 (0.93)	3.48 (0.86)	3.4 (1.01)	3.38 (0.90)
Fearful/Insecure	3.06 (0.81)	3.26 (0.81)	3.03 (1.10)	3.12 (1.06)
Intelligent	4.83 (0.84)	4.76 (0.77)	4.58 (0.83)	4.61 (0.76)
Negative Affect	3.45 (1.08)	3.51 (1.03)	2.99 (1.07)	3.0 (1.03)
Open to Experience	4.99 (0.82)	5.12 (0.79)	4.85 (0.88)	4.98 (0.79)
Organized	4.76 (0.81)	4.88 (0.79)	4.79 (0.95)	4.89 (0.84)
Positive Emotions	4.86 (0.91)	5.05 (0.93)	4.74 (0.98)	4.96 (0.90)
Shy	2.97 (0.87)	2.9 (0.93)	3.18 (0.94)	3.0 (1.04)
Sociable	4.89 (0.95)	4.9 (0.98)	4.58 (0.93)	4.68 (0.93)
Strong Willed	4.15 (0.82)	4.06 (0.79)	3.92 (0.75)	4.05 (0.91)
Neuroticism	3.25 (0.83)	3.39 (0.80)	3.01 (0.96)	3.06 (0.92)
Extraversion	4.89 (0.70)	4.98 (0.74)	4.76 (0.75)	4.87 (0.68)
Openness	4.83 (0.84)	4.76 (0.77)	4.58 (0.83)	4.61 (0.76)
Disagreeableness	3.5 (0.58)	3.37 (0.54)	3.39 (0.62)	3.38 (0.66)
Conscientiousness	4.76 (0.73)	4.78 (0.73)	4.69 (0.83)	4.72 (0.75)

report). Extraversion was defined by sociability, activity, and positive emotionality; Disagreeableness was characterized by antagonism and strong will; Neuroticism

captured fear/insecurity, negative affectivity, and shyness; Conscientiousness was described by organized behavior, achievement orientation and low distractibility; and Openness comprised intelligence and openness to experience. The Compliant and Considerate scales were omitted from the model due to their low discriminative validity (Knyazev et al., 2008). However, they were used in the analysis of lower-order traits.

Gender differences in each country, age group, and with respect to the method of assessment in early adolescence were calculated by *t*-test. Only the differences at 0.05 probability level or lower were accounted for. The magnitude of differences (effect size) was defined by *d*, a measure of how many standard deviations apart the two means are: a *d* of 0.2, 0.5, and 0.8 represents a small, medium, and large effect size, respectively (Cohen, 1988).

In the two countries, the direction of gender differences was quite similar for one informant, but differed between the informants. The differences between boys and girls in Slovene and Russian parent ratings of the total sample and in the Slovene and Russian adolescent self-ratings strongly correlated with each other (rank-order $r = .94$ and $.76$ for parent and self-ratings, respectively, both $ps < .001$). Contrariwise, within-country correlations of *d*s between parent and self-ratings of adolescents were substantially weaker ($r = .47$, $p = .038$ and $r = .36$, $p = .114$ for the Russian and the Slovene sample, respectively).

Gender differences in parent ratings of child trait expression across age groups

The effect sizes of gender differences derived from parent ratings across the four age groups and across the two countries are captured in Table 2. The magnitude of gender differences is small and similar in the two countries. Mean *d*s across the 20 variables were 0.16 in the Russian and 0.14 in the Slovene sample. Overall, in both countries, girls were rated by their parents as somewhat more achievement orientated, compliant, considerate, and organized, whereas boys were assessed as slightly more active, antagonistic, and distractible. Additionally, in Slovenia, boys were perceived as slightly more fearful/insecure, shy, and prone to negative affect, whereas in Russia, they were perceived as less intelligent than girls. On the higher-order scales, according to parent ratings, girls appeared more conscientious and agreeable in both countries. In Slovenia, girls were also assessed as slightly more emotionally stable, and in Russia as more open than boys. However, these descriptions present general trends in parental perceptions of their children's personality trait expression by child gender as all of the gender differences across the four age groups did not reach statistical significance.

Age trends in gender differences

As indicated in Table 2, gender differences across the two countries appear significant with regard to more personality traits in school age children and early

Table 2. Mean z-score gender differences (d) according to parent ratings of four age groups (2–3, 4–6, 7–11, and 12–15 years) in the two countries.

ICID Scale	2–3 years		4–6 years		7–11 years		12–15 years	
	Slovenia	Russia	Slovenia	Russia	Slovenia	Russia	Slovenia	Russia
Achievement Orientation	-0.21*	-0.06	-0.08	-0.15	-0.03	-0.22**	-0.42***	-0.39***
Activity Level	0.11*	0.23	0.05	0.36**	0.28**	0.19*	0.32***	0.12
Antagonism	0.20*	-0.02	0.20	0.01	0.05	0.12	0.35***	0.19*
Compliant	-0.19	-0.15	-0.06	-0.08	-0.03	-0.15	-0.32***	-0.25**
Considerate	-0.27**	-0.23	-0.13	-0.02	-0.08	-0.13	-0.18	-0.18*
Distraction	0.25*	0.36	0.02	0.09	0.20*	0.29***	0.20*	0.37***
Fearful/Insecure	0.14	-0.07	0.27*	-0.13	0.09	0.12	0.10	0.06
Intelligent	-0.27**	-0.19	-0.13	-0.13	0.14	-0.25***	-0.11	-0.32***
Negative Affect	0.04	0.04	0.13	0.02	0.17*	0.13	0.12	0.02
Open to Experience	-0.22*	0.14	0.00	0.13	0.17*	-0.09	-0.08	-0.07
Organized	-0.20*	-0.15	-0.06	-0.20	-0.21**	-0.26***	-0.52***	-0.45***
Positive Emotions	-0.25*	-0.10	0.03	0.04	0.01	0.00	-0.11	-0.20*
Shy	0.26**	-0.09	0.18	-0.13	0.14*	0.05	0.18	0.13
Sociable	-0.21*	-0.05	0.01	0.11	-0.06	0.02	-0.10	-0.13
Strong Willed	-0.09	0.16	-0.03	0.06	0.02	0.14	-0.05	-0.17
Neuroticism	0.10	-0.01	0.22	-0.05	0.15*	0.15	0.13	0.05
Extraversion	-0.22*	0.02	-0.04	0.16	-0.04	0.08	-0.06	-0.12
Openness	-0.27**	-0.19	-0.13	-0.13	-0.07	0.14	-0.11	-0.32***
Disagreeableness	0.16	0.12	0.09	0.06	0.17*	0.10	0.29***	0.14
Conscientiousness	-0.26*	-0.22	-0.06	-0.18	-0.26**	-0.17*	-0.42***	-0.45***

Note. Positive value indicates that boys scored higher than girls.

* $p < .05$; ** $p < .01$; *** $p < .001$.

adolescents relative to preschoolers whereas many small, though significant gender differences were also demonstrated in parental ratings of Slovene toddlers (note that the Russian sample of parents reporting on toddlers was considerably smaller). Actually, the mean *ds* were higher in the Slovene parents' assessments of toddlers (0.20) and early adolescents (0.21) than in ratings of preschoolers (0.10) and school children (0.12). The age trend in mean effect sizes of gender differences was 0.13, 0.11, 0.14, and 0.21 for assessments of Russian toddlers, preschoolers, school children, and early adolescents, respectively. A relatively consistent pattern of significant gender differences across countries was identified in parental reports on the two older groups: school age and adolescent daughters were ascribed higher levels of conscientiousness, organized behavior, and achievement orientation than were sons of the same age who were, in turn, assessed as more active and distractible in comparison to daughters.

Gender differences in self- vs. parent ratings of personality

Table 3 provides mean effect sizes of gender differences in lower- and higher-order personality traits across countries (Slovenia vs. Russia), and methods of assessment (parent vs. self-reports). The *ds* for gender differences according to parental data source are presented with regard to the total sample of target children/adolescents and the column *Self* provides *ds* for gender differences in early adolescents. Considering the adolescent age group only, the *ds* for self-report in Table 3 are compared with *ds* derived from parental ratings of the same age group displayed across the last two columns of the Table 2.

According to adolescent self-ratings, girls in both countries appeared significantly more considerate and prone to positive emotions than boys, who in turn, tended to rate themselves as more antagonistic (significantly in Slovenia) and active (significantly in Russia). Additionally, the Russian girls appeared to be significantly less shy in comparison to Russian boys, whereas Slovene girls provided significantly higher ratings with regard to fear/insecurity than did their male counterparts. On the higher-order scales, Slovene boys rated themselves as more disagreeable than Slovene girls.

As with overall parent reports on children/adolescents (mean *ds* for Slovenia and Russia were 0.14 and 0.16, respectively), and specifically on adolescents (mean *ds* were 0.21 in each country), the mean magnitude of gender differences in adolescent self-ratings were small: 0.14 in the Slovene and 0.10 in the Russian sample. Even the largest effect size did not exceed a difference of 0.39 *SD* in mean scores for self-reports on the Considerate scale (Slovene girls scored higher than Slovene boys). Except for a 0.52 *SD* difference (a medium effect size) in parent reports on their adolescent daughters' and sons' organized behavior (Slovene girls were rated higher than Slovene boys), all other gender differences derived from parental data on early adolescents were small in both countries. However, analyses of parental

versus self-ratings yielded significant gender differences in perceived expression of more adolescent personality traits as well as somewhat larger effect sizes of these differences. Several similarities in the pattern of significant gender differences relying on parental data sources were obtained across Slovenia and Russia: parents rated adolescent-girls as more achievement orientated, compliant, organized and conscientious than adolescent-boys who were, in turn, assessed as more antagonistic and distractible than were girls.

Table 3. Mean z-score difference (*d*) between adolescent boys and girls according to overall parent- and adolescent self-ratings.

	Russia		Slovenia	
	Parent	Self	Parent	Self
Achievement Orientation	-0.28***	0.05	-0.19***	0.08
Activity Level	0.22***	0.19*	0.22***	0.19
Antagonism	0.13*	0.17	0.18***	0.35***
Compliant	-0.20***	0.03	-0.16***	-0.00
Considerate	-0.14**	-0.22*	-0.16***	-0.39***
Distractible	0.30***	0.03	0.19***	0.01
Fearful/Insecure	0.06	-0.08	0.15**	-0.26*
Intelligent	-0.26***	-0.03	-0.07	0.08
Negative Affect	0.08	-0.01	0.13**	-0.06
Open to Experience	-0.02	-0.16	0.00	-0.16
Organized	-0.32***	-0.11	-0.28***	-0.15
Positive Emotions	-0.10	-0.23**	-0.08	-0.20*
Shy	0.05	0.18*	0.14**	0.08
Sociable	-0.01	-0.11	-0.06	-0.00
Strong Willed	-0.04	-0.16	0.02	0.12
Neuroticism	0.07	-0.06	0.16***	-0.16
Extraversion	-0.02	-0.15	-0.04	-0.12
Openness	-0.26***	-0.04	-0.07	0.08
Disagreeableness	0.13**	0.01	0.17***	0.23*
Conscientiousness	-0.34***	-0.04	-0.25***	-0.02

Note. Positive value indicates that boys scored higher than girls.

* $p < .05$; ** $p < .01$; *** $p < .001$.

Overall, in both countries the pattern of gender differences in self- and parent reports was not similar. At the higher-order level, only higher disagreeableness of Slovene boys relative to Slovene girls was confirmed by both raters (also when specifically accounting for parent reports on adolescents). At the lower-order traits, significantly higher consideration of girls in comparison to boys was obtained across

countries and methods of assessment, the two Slovene groups of informants concurred in rating boys significantly higher for antagonism than girls (also when parents reported specifically on adolescents), whereas significantly higher activity level of boys relative to girls across the two data sources was demonstrated in Russia.

Discussion

Child and adolescent personality in two different cultural contexts (Slovenia and Russia) was assessed by parents and adolescents themselves. The higher- and lower-order personality traits were recorded through a specially designed assessment tool to capture child and early adolescent characteristics that are the most salient for parents across various cultural settings, the ICID (Halverson et al., 2003). Gender differences in the ICID trait scores were studied and compared between the two countries, across age groups and methods of assessment (parent- and self-report). As expected, even the significant gender differences were small across countries, age of the target children/adolescents, and informants reporting on adolescent trait expression. Gender differences as reflected from parental reports on children were present even among toddlers but appeared significant in Slovenia only, partly due to a small Russian sample size. However, the small differences between parents' ratings of toddler-boys and toddler-girls in both countries dropped in early childhood and reappeared significant with school age children, while the magnitude of gender differences increased in early adolescence. This pattern thus, does not lend unambiguous evidence in support of biosocial hypothesis.

According to our prediction, the magnitude of gender differences was similar in Slovenia and Russia but patterns of these differences were substantially different across the raters. The between-informants difference seems to be robust since it generalized across the two countries. The observed patterns of gender differences were much closer to each other for one source of information across the countries than they were for different sources of information within the country. It has been noted previously with adult samples that external ratings, as compared to self-ratings, systematically underestimate the magnitude of age and gender differences (McCrae, Costa, Hřebičková et al., 2004; McCrae, Costa, Martin et al., 2004). However, this was not the case in our study as the overall magnitude of personality gender differences was similar across parent- and self-reports, while the parental assessments yielded significant gender differences in more traits than the adolescent self-ratings. Considering the magnitude of gender differences derived from parental reports on adolescents only, the effect sizes were even higher than those demonstrated for self-ratings.

The direction of personality gender differences based on parent ratings appeared consistent to findings of other studies (e.g., Halverson, 2003; Kohnstamm, 1989; Zupančič et al., 2006; Zupančič & Kavčič, 2005, 2007), indicating that girls

are considered by their parents as slightly more expressive in socially desirable traits than boys across different cultural environments. Parents rated girls somewhat higher than boys on Conscientiousness (and on the respective lower-order traits) although the differences generally reached significance in school age children/adolescents. This lends support to previous findings on gender effect for Conscientiousness in free descriptive research (De Fruyt et al., 1998), parental ICID ratings (Halverson, 2003; Slobodskaya, 2005; Zupančič et al., 2006), and temperament assessments of children's effortful control (Else-Quest et al., 2006). Furthermore, the gender differences resulting from parental reports on adolescents' conscientiousness and organized behavior were larger in both countries than the differences obtained with assessments of children but no respective gender effect was established in adolescent self-ratings. The small effect of gender on Disagreeableness and on the Considerate scale (resulting also from self-reports) is in line with studies demonstrating that females score higher in the Agreeableness domain (Costa et al., 2001; McCrae, Terraciano et al., 2005), although other studies established no differences between boys and girls on this robust trait (De Fruyt et al., 1998; Zupančič et al., 2006). In addition, a higher level of activity in boys found in our study was in accord with findings from the temperament research (Else-Quest et al., 2006; Kohnstamm, 1989) and the results obtained from free parental descriptions (De Fruit et al., 1998). The gender differences in child personality trait expression may actually exist but there could be also a specific gender stereotypic bias in parental perception of their daughters' and sons' behavior as the differences between boys and girls derived from adolescents' self-reports did not appear in the same direction for several traits or were even absent. Alternatively, the gender differences in child/adolescent behavior may be exhibited differently at home (where parents observe it and create child personality judgments) than outside home, e.g. among the class-mates (where the adolescents provided self-report).

Inconsistently to the well-documented findings of females being less emotionally stable than males obtained through observer and self-reported personality questionnaires (e.g., McCrae, Terraciano et al., 2005), but in accord with results based on free parental descriptions (De Fruyt et al., 1998), the gender differences in the Neuroticism domain were mostly non-significant with our data. On the contrary, when the differences reached significance they appeared in the opposite direction. Overall, the Slovene parents, for example, rated boys higher on Neuroticism (and on all of the three corresponding lower-order traits) than girls, and the Russian adolescent boys reported on higher levels of shyness than their female peers. Our hypothesis that the personality scores would reflect more emotional instability in girls at least in adolescence was supported only with respect to the Fearful/Insecure facet of Neuroticism and specifically with Slovene self-reports. Another unexpected finding of our study was Russian parents' tendency to rate their daughters higher than sons on the Openness domain. This was mostly due to higher assessments of girls on the Intelligent scale, with gender differences reaching significance for the

two older groups of Russian children/adolescents as well as for Slovene toddlers. The results regarding both gender differences in Neuroticism and Openness/Intelligence are in contrast to the findings with free parental descriptions of children from seven countries (De Fruit et al., 1998) and to those one might have expected according to traditional stereotypes about gender roles. However, similarly to the interpretation by Costa et al. (2001), it might be equally plausible that parents more often emphasize intelligence and emotional stability in daughters than in sons because boys are expected to be bright and stable according to their gender role, while girls' intelligence and emotional stability is attributed to their personality.

Conclusions

The study compared gender differences in child/adolescent personality trait expression across two European countries. In both Slovenia and Russia large samples of target individuals were assessed by their caregivers and adolescents provided self-reports employing an internationally recognized age- and culture-neutral personality measure. To our knowledge no cross-cultural studies of child personality gender differences in mean level trait expression using the ICID data were published so far.

Overall, small gender effects in our study were in accord with expectations and their direction was similar across countries for each source of information. Notable dissimilarities between the informants in both countries were demonstrated, so that higher Agreeableness and Conscientiousness of females were observed for parent reports only, while the results for Neuroticism were inconclusive. Although gender differences in many traits were evident from parental reports on toddlers, the effect sizes did not increase with age of the targets until they reached adolescence. Both the absence of a clear pattern in magnitude of gender differences over age and a cross-sectional design of our study preclude an interpretation of gender differences as a result of personality development. Our findings indicate that, along with biological and social influences that presumably shape personality by gender, parental response bias may also contribute to the perceived gender differences in child/adolescent personality expression. In addition, there could be cohort differences in both gender differences in child/adolescent personality and in informants' response bias.

There are a number of limitations of this study one of which is that it was cross-sectional, and although the data came from different informants, they did not rate the same individuals altogether. Future studies should address this shortcoming using longitudinal and multiple-informant data. It may also be assumed that the Slovene sample was more representative for the population than the Russian sample which was actually western Siberian. Further, our results refer to parent evaluations of child/adolescent personality and adolescent self-reports. These do not merely reflect objectively recorded characteristics, but also parental and adolescent implicit theories on child and own personality, respectively. The field would therefore benefit

from intra- and inter-cultural studies on gender differences in developing personality by comparing observation measures with adult- and adolescent self-assessed child/adolescent inventories.

References

- Allik, J., Laidra, K., Realo, A., & Pullman, H. (2004). Personality development from 12 to 18 years of age: Changes in mean levels and structure of traits. *European Journal of Personality, 18*, 445-462.
- Barbaranelli, C., Caprara, G. V., Rabasca, A., & Pastorelli, C. (2003). A questionnaire for measuring the Big Five in late childhood. *Personality and Individual Differences, 34*, 645-664.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Costa, P. T., Jr., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology, 81*, 322-331.
- De Fruyt, F., Van Hiel, A., & Buyst, V. (1998). Parental personality descriptors of boys and girls. In G. A. Kohnstamm, C. F. Halverson, Jr., I. Mervielde, & V. L. Havill (Eds.), *Parental descriptions of child personality: Developmental antecedents of the Big Five?* (pp. 155-167). Hillsdale, NJ: Erlbaum.
- Donahue, E. M. (1994). Do children use the Big Five, too? Content and structural form in personality description. *Journal of Personality, 62*(1), 45-66.
- Else-Quest, N. M., Hyde, J. S., Goldsmith, H. H., & Van Hulle, C. A. (2006). Gender differences in temperament: A meta-analysis. *Psychological Bulletin, 132*, 33-72.
- Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin, 116*, 429-456.
- Gartstein, M. A., Slobodskaya, H. R., & Kinsht, I. A. (2003). Cross-cultural differences in the first year of life: United States of America (U.S.) and Russian. *International Journal of Behavioral Development, 27*, 316-328.
- Goldberg, L. R. (2001). Analyses of Digman's child-personality data: Derivation of Big-Five factor scores from each of six samples. *Journal of Personality, 69*(5), 709-743.
- Halverson, C. F. (2003). *Culture, age, and personality development*. Paper presented at the 11th European Conference on Developmental Psychology, Milan, Italy.
- Halverson, C. F., Havill, V. L., Deal, J., Baker, S. R., Victor, J. B., Pavlopoulos, V., Besevegis, E., & Wen, L. (2003). Personality structure as derived from parental ratings of free descriptions of children: The Inventory of Child Individual Differences. *Journal of Personality, 71*(6), 995-1026.
- Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviours, institutions, and organizations across nations* (2nd ed.). Thousand Oaks, CA: Sage.
- Kavčič, T., & Zupančič, M. (2006). *Osebnost otrok in njihovi medsebojni odnosi v družini [Child personality and interpersonal relationships within families]*. Ljubljana: Znanstvenoraziskovalni inštitut Filozofske fakultete.
- Knyazev, G. G., & Slobodskaya, H. R. (2005). Five factor personality structure in children and adolescents (based on parent and self-reports). *Psychological Journal,*

26(6), 59-67 (In Russian).

- Knyazev, G. G., Zupančič, M., & Slobodskaya, H. R. (2008). Child personality in Slovenia and Russia. Structure and mean level of traits in parent and self-ratings. *Journal of Cross-Cultural Psychology*, 39, 317-334.
- Kohnstamm, G. A. (1989). Temperament in childhood: Cross-cultural and sex differences. In G. A. Kohnstamm, J. E. Bates & M. K. Rothbart (Eds), *Temperament in childhood*, (pp. 483-508). New York: Wiley.
- Lynn, R., & Martin, T. (1997). Gender differences in Extraversion, Neuroticism, and Psychoticism in 37 nations. *Journal of Social Psychology*, 137, 369-373.
- Maccoby, E., E., & Jacklin, C. N. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University Press.
- Markey, P. M., Markey, C. N., Ericksen, A. J., & Tinsley, B. J. (2002). A preliminary validation of preadolescents' self-reports using the Five-Factor Model of personality. *Journal of Research in Personality*, 36(2), 173-181.
- McCrae, R. R., & Costa, P. T. Jr. (2003). *Personality in adulthood: A Five-Factor Theory perspective* (2nd ed.). New York: Guilford.
- McCrae, R. R., Costa, P. T. Jr., Hřebíčková, M., Urbánek, T., Martin, T. A., Oryol, V. E., Rukavishnikov, A. A., & Senin, I. G. (2004). Age differences in personality traits across cultures: Self-report and observer perspectives. *European Journal of Personality*, 18, 143-157.
- McCrae, R. R., Costa, P. T. Jr., Martin, T. A., Oryol, V. E., Rukavishnikov, A. A., Senin, I. G., Hřebíčková, M., & Urbánek, T. (2004). Consensual validation of personality traits across cultures. *Journal of Research in Personality*, 38, 179-201.
- McCrae, R. R., Costa, P. T. Jr., Ostendorf, F., Angleitner, A., Hřebíčková, M., Avia, M. D., Sanz, J., Sanchez-Bernardos, M. L., Kusdil, M. E., Woodfield, R., Saunders, P. R., & Smith, P. B. (2000). Nature over nurture: Temperament, personality and lifespan development. *Journal of Personality and Social Psychology*, 78, 173-186.
- McCrae, R. R., Costa, P. T. Jr., Terraciano, A., Parker, W. D., Mills, C. J., De Fruyt, F., & Mervielde, I. (2002). Personality trait development from age 12 to age 18: Longitudinal, cross-sectional, and cross-cultural analyses. *Journal of Personality and Social Psychology*, 83, 1456-1468.
- McCrae, R. R., Martin, T. A., & Costa, P. T. Jr. (2005). Age trends and age norms for the NEO Personality Inventory-3 in adolescents and adults. *Assessment*, 12, 363-373.
- McCrae, R. R., Terraciano, A., & 78 members of the Personality Profiles Project (2005). Universal features of personality traits from the observer's perspective: Data from 50 cultures. *Journal of Personality and Social Psychology*, 88, 547-561.
- Measelle, J. R., John, O. P., Ablow, J. C., Cowan, P. A., & Cowan, C. P. (2005). Can children provide coherent, stable, and valid self-reports on the Big Five dimensions? A longitudinal study from ages 5 to 7. *Journal of Personality and Social Psychology*, 89, 90-106.
- Mervielde, I., Buyst, V., & De Fruyt, F. (1995). The validity of the Big-Five as a model for teachers' ratings of individual differences among children aged 4-12 years. *Personality and Individual Differences*, 18(4), 525-534.
- Mervielde, I. & De Fruyt, F. (1999). Construction of the Hierarchical Personality Inventory for Children (HiPIC). In I. Mervielde, I. Deary, F. De Fruyt & F. Ostendorf (Eds.), *Personality psychology in Europe: Vol. 7* (pp. 107-128). Tilburg: Tilburg University Press.

- Resing, W. C. M., Bleichrodt, N., & Dekker, P. H. (1999). Measuring personality traits in the classroom. *European Journal of Personality, 13*, 493-509.
- Scholte, R. H. J., van Aken, M. A. G., & van Lieshout, C. F. M. (1997). Adolescent personality factors in self-ratings and peer nominations and their prediction of peer acceptance and peer rejection. *Journal of Personality Assessment, 69*(3), 534-554.
- Shiner, R. L., & Caspi, A. (2003). Personality differences in childhood and adolescence: Measurement, development, and consequences. *Journal of Child Psychology and Psychiatry, 44*(1), 2-32.
- Slobodskaya, H. R. (2005). *Personality as predictor of behavioural and emotional problems in Russian children at different ages*. Paper presented at the 12th European Conference on Developmental Psychology, La Laguna, Spain.
- Slobodskaya, H. R., Safronova, M. V., Knyazev, G. G., & Wilson, G. D. (2001). Reactions of Russian adolescents to reward and punishment: A cross-cultural study of the Gray-Wilson Personality Questionnaire. *Personality and Individual Differences, 30*, 1211-1224.
- Van Lieshout, C. F. M., & Haselager, G. J. T. (1994). The Big Five personality factors in Q-sort descriptions of children and adolescents. In C. F. Halverson, Jr., G. A. Kohnstamm & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 293-318). Hillsdale, NJ: Lawrence Erlbaum.
- World Bank (2004). *World development indicators*. Washington, DC: World Bank Publications.
- Zupančič, M., Gril, A., & Kavčič, T. (2006). Child and adolescent personality: Its structure, age trends and gender differences. *Studia Psychologica, 48*(4), 311-332.
- Zupančič, M., & Kavčič, T. (2005). Gender differences in personality through early childhood: A multiple-informant perspective. *Psihološka obzorja [Horizons of Psychology], 14*(2), 11-38.
- Zupančič, M., & Kavčič, T. (2007). *Otroci od vrtca do šole: razvoj osebnosti in socialnega vedenja ter učna uspešnost prvošolcev [The children from pre-school to school: Development of personality and social behaviour, and academic achievement in first-graders]*. Ljubljana: Znanstvenoraziskovalni inštitut Filozofske fakultete.
- Zupančič, M., Sočan, G., & Kavčič, T. (2007). Consistency in adult reports on child personality over the pre-school years. *European Journal of Developmental Psychology*. Republished on August 18, 2007, DOI: 10.1080/17405620701439887.

Appendix

Table A1. Means (and standard deviations in parentheses) in ICID parent ratings: Boys and girls across four age groups (2–3, 4–6, 7–11, and 12–15 years) – Slovene sample ($N = 1850$).

ICID Scale	Total Sample		2–3 years		4–6 years		7–11 years		12–15 years	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Achievement	4.73 (0.84)	4.89 (0.86)	4.66 (0.70)	4.8 (0.64)	4.75 (0.79)	4.82 (0.86)	4.77 (0.92)	4.79 (0.89)	4.74 (0.88)	5.13 (0.93)
Activity Level	5.1 (0.93)	4.89 (0.98)	5.11 (0.90)	5.02 (0.82)	5.13 (0.89)	5.07 (1.26)	5.13 (0.97)	4.87 (0.94)	5.01 (0.92)	4.7 (0.94)
Antagonism	2.89 (0.86)	2.73 (0.92)	3.17 (0.84)	3.0 (0.76)	3.03 (0.82)	2.86 (0.92)	2.73 (0.86)	2.68 (1.03)	2.81 (0.85)	2.51 (0.85)
Compliant	4.63 (0.81)	4.76 (0.79)	4.39 (0.83)	4.52 (0.61)	4.64 (0.73)	4.68 (0.79)	4.68 (0.82)	4.71 (0.79)	4.79 (0.80)	5.05 (0.82)
Considerate	5.09 (0.82)	5.23 (0.87)	4.92 (0.67)	5.1 (0.69)	5.07 (0.76)	5.17 (0.81)	5.13 (0.85)	5.2 (0.94)	5.23 (0.87)	5.39 (0.92)
Distractible	3.59 (0.83)	3.43 (0.81)	3.60 (0.65)	3.43 (0.71)	3.49 (0.75)	3.47 (0.83)	3.67 (0.87)	3.5 (0.87)	3.49 (0.96)	3.32 (0.81)
Fearful/Insecure	3.55 (0.83)	3.43 (0.81)	3.58 (0.68)	3.48 (0.76)	3.64 (0.79)	3.42 (0.80)	3.58 (0.89)	3.5 (0.83)	3.41 (0.88)	3.33 (0.85)
Intelligent	4.98 (0.84)	5.03 (0.81)	4.85 (0.71)	5.04 (0.74)	5.07 (0.81)	5.17 (0.82)	5.05 (0.88)	4.94 (0.82)	4.96 (0.90)	5.06 (0.80)
Negative Affect	3.7 (0.98)	3.57 (1.00)	3.75 (0.86)	3.71 (0.88)	3.77 (0.88)	3.64 (1.06)	3.72 (1.03)	3.54 (1.01)	3.58 (1.04)	3.45 (1.04)
Open to Experience	5.08 (0.79)	5.08 (0.78)	4.99 (0.75)	5.16 (0.75)	5.20 (0.75)	5.2 (0.80)	5.14 (0.80)	5.0 (0.77)	4.99 (0.80)	5.06 (0.80)
Organized	4.44 (0.79)	4.67 (0.83)	4.46 (0.64)	4.58 (0.62)	4.51 (0.77)	4.56 (0.86)	4.4 (0.83)	4.59 (0.87)	4.43 (0.85)	4.89 (0.87)
Positive Emotions	5.11 (0.81)	5.17 (0.81)	4.98 (0.72)	5.16 (0.71)	5.14 (0.89)	5.12 (0.79)	5.11 (0.80)	5.11 (0.86)	5.18 (0.81)	5.28 (0.84)
Shy	3.2 (0.86)	3.08 (0.90)	3.27 (0.77)	3.05 (0.88)	3.22 (0.78)	3.07 (0.93)	3.12 (0.92)	3.08 (0.88)	3.27 (0.89)	3.12 (0.92)
Sociable	4.85 (0.85)	4.9 (0.83)	4.70 (0.77)	4.86 (0.82)	4.90 (0.77)	4.89 (0.86)	4.93 (0.90)	4.91 (0.79)	4.93 (0.90)	4.91 (0.86)
Strong Willed	4.1 (0.80)	4.08 (0.81)	4.29 (0.77)	4.36 (0.80)	4.19 (0.87)	4.22 (0.91)	4.08 (0.75)	3.97 (0.76)	4.08 (0.75)	3.92 (0.75)
Neuroticism	3.63 (0.79)	3.5 (0.81)	3.66 (0.68)	3.6 (0.70)	3.71 (0.74)	3.53 (0.82)	3.65 (0.82)	3.52 (0.82)	3.65 (0.82)	3.39 (0.86)
Extraversion	5.01 (0.64)	5.03 (0.67)	4.91 (0.62)	5.04 (0.62)	5.04 (0.57)	5.07 (0.71)	5.05 (0.67)	5.0 (0.65)	5.05 (0.67)	5.04 (0.70)
Openness	4.98 (0.84)	5.03 (0.81)	4.85 (0.71)	5.04 (0.74)	5.07 (0.81)	5.17 (0.82)	5.05 (0.88)	4.94 (0.82)	5.05 (0.88)	5.06 (0.85)
Disagreeableness	3.46 (0.62)	3.35 (0.65)	3.7 (0.55)	3.61 (0.55)	3.53 (0.62)	3.47 (0.71)	3.38 (0.62)	3.31 (0.65)	3.38 (0.62)	3.13 (0.60)
Conscientiousness	4.53 (0.72)	4.71 (0.73)	4.51 (0.58)	4.65 (0.54)	4.59 (0.69)	4.64 (0.76)	4.5 (0.76)	4.63 (0.76)	4.5 (0.76)	4.9 (0.76)

Table A2. Means (and standard deviations in parentheses) in ICID parent ratings: Boys and girls across four age groups (2–3, 4–6, 7–11, and 12–15 years) – Russian sample ($N = 1635$).

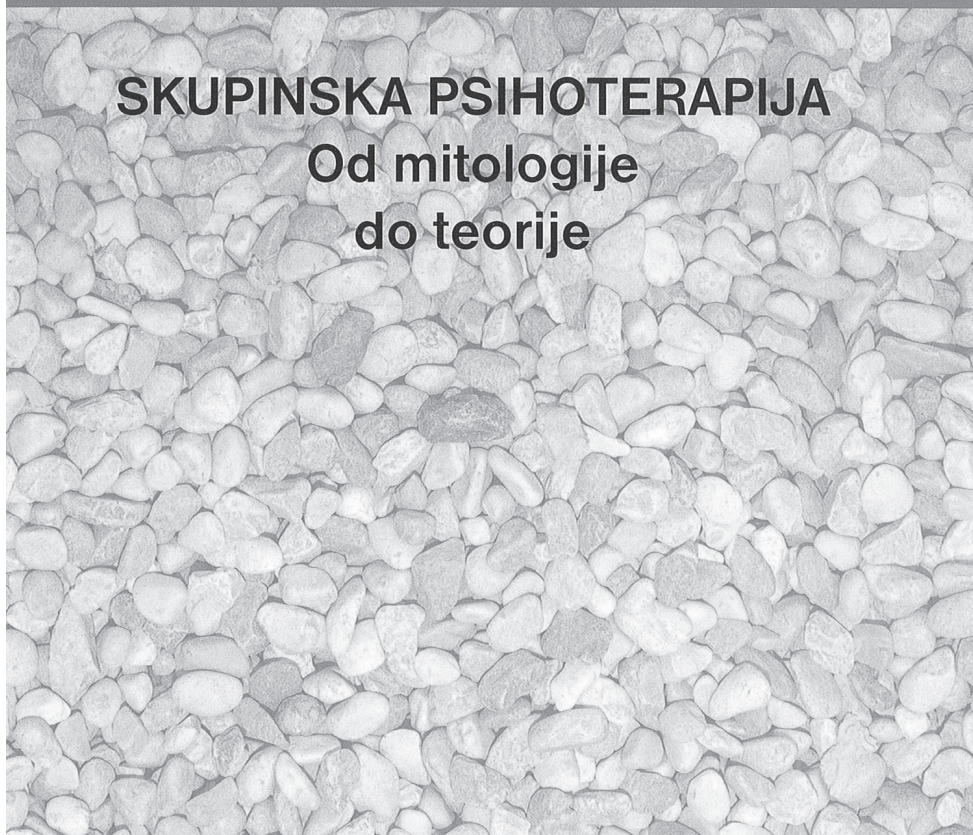
ICID Scale	Total Sample		2–3 years		4–6 years		7–11 years		12–15 years	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Achievement	4.32 (0.89)	4.57 (0.91)	4.43 (0.80)	4.47 (0.67)	4.28 (0.77)	4.41 (0.88)	4.32 (0.84)	4.51 (0.87)	4.33 (1.06)	4.73 (0.97)
Activity Level	4.73 (0.96)	4.52 (0.94)	4.85 (0.80)	4.67 (0.84)	5.00 (0.88)	4.68 (0.84)	4.74 (0.91)	4.57 (0.93)	4.50 (1.07)	4.37 (0.98)
Antagonism	2.84 (0.97)	2.71 (0.98)	2.84 (0.89)	2.86 (1.07)	2.94 (0.95)	2.93 (0.97)	2.85 (0.96)	2.73 (1.03)	2.76 (1.01)	2.58 (0.91)
Compliant	4.38 (0.79)	4.54 (0.80)	4.18 (0.78)	4.29 (0.61)	4.29 (0.67)	4.35 (0.71)	4.34 (0.75)	4.4 (0.80)	4.55 (0.90)	4.76 (0.81)
Considerate	4.81 (0.89)	4.94 (0.86)	4.82 (0.74)	5.01 (0.90)	4.82 (0.85)	4.84 (0.87)	4.78 (0.87)	4.9 (0.86)	4.86 (0.97)	5.01 (0.85)
Distractible	3.87 (0.84)	3.62 (0.84)	3.74 (0.71)	3.48 (0.70)	3.78 (0.75)	3.71 (0.73)	3.98 (0.78)	3.74 (0.85)	3.83 (0.98)	3.48 (0.87)
Fearful/Insecure	3.79 (0.88)	3.75 (0.85)	3.60 (0.66)	3.66 (0.92)	3.76 (0.88)	3.84 (0.81)	3.92 (0.82)	3.82 (0.80)	3.69 (0.98)	3.64 (0.88)
Intelligent	4.4 (0.82)	4.61 (0.78)	4.6 (0.85)	4.74 (0.64)	4.53 (0.78)	4.63 (0.75)	4.3 (0.74)	4.5 (0.77)	4.41 (0.93)	4.69 (0.82)
Negative Affect	3.28 (1.0)	3.19 (0.99)	3.24 (0.88)	3.2 (1.01)	3.33 (0.97)	3.31 (0.99)	3.38 (0.99)	3.24 (1.04)	3.12 (1.04)	3.09 (0.95)
Open to Experience	4.84 (0.78)	4.86 (0.80)	5.04 (0.78)	4.93 (0.79)	5.00 (0.78)	4.9 (0.79)	4.76 (0.77)	4.83 (0.82)	4.79 (0.78)	4.85 (0.78)
Organized	4.26 (0.85)	4.54 (0.86)	4.44 (0.86)	4.56 (0.78)	4.23 (0.80)	4.36 (0.81)	4.23 (0.81)	4.44 (0.82)	4.29 (0.93)	4.71 (0.91)
Positive Emotions	4.93 (0.87)	5.02 (0.83)	5.11 (0.79)	5.19 (0.81)	5.10 (0.87)	5.06 (0.83)	4.89 (0.87)	4.97 (0.86)	4.85 (0.87)	5.01 (0.80)
Shy	3.28 (0.96)	3.23 (0.94)	3.03 (0.79)	3.11 (1.04)	3.07 (0.97)	3.19 (0.88)	3.33 (0.92)	3.22 (0.90)	3.41 (1.02)	3.29 (0.99)
Sociable	4.5 (0.86)	4.51 (0.85)	4.44 (0.69)	4.47 (0.79)	4.70 (0.88)	4.6 (0.86)	4.46 (0.79)	4.45 (0.80)	4.44 (0.96)	4.54 (0.90)
Strong Willed	3.99 (0.69)	4.02 (0.74)	4.27 (0.70)	4.15 (0.74)	4.16 (0.68)	4.12 (0.74)	3.93 (0.70)	3.95 (0.69)	3.9 (0.65)	4.02 (0.78)
Neuroticism	3.54 (0.84)	3.54 (0.82)	3.47 (0.82)	3.43 (0.90)	3.55 (0.84)	3.59 (0.83)	3.65 (0.78)	3.54 (0.82)	3.41 (0.93)	3.36 (0.80)
Extraversion	4.75 (0.67)	4.75 (0.67)	4.77 (0.68)	4.86 (0.69)	4.92 (0.69)	4.82 (0.68)	4.71 (0.63)	4.75 (0.67)	4.67 (0.72)	4.75 (0.68)
Openness	4.4 (0.82)	4.50 (0.77)	4.61 (0.78)	4.74 (0.64)	4.53 (0.78)	4.63 (0.75)	4.3 (0.74)	4.50 (0.77)	4.41 (0.93)	4.69 (0.82)
Disagreeableness	3.49 (0.62)	3.43 (0.64)	3.41 (0.63)	3.57 (0.60)	3.61 (0.58)	3.57 (0.59)	3.49 (0.61)	3.43 (0.64)	3.37 (0.65)	3.28 (0.62)
Conscientiousness	4.23 (0.75)	4.4 (0.74)	4.5 (0.77)	4.51 (0.64)	4.25 (0.66)	4.37 (0.70)	4.19 (0.69)	4.4 (0.74)	4.26 (0.90)	4.66 (0.81)



Univerza v Ljubljani
FILOZOFSKA
FAKULTETA

Peter Praper

SKUPINSKA PSIHOTERAPIJA
Od mitologije
do teorije



From chaos through cosmos toward coinonia: A group-analytic developmental line

*Peter Praper**

University of Ljubljana, Department of Psychology, Ljubljana, Slovenia

Abstract: The phenomenon of group cohesion is among most examined in group psychotherapy and repeatedly confirmed as the one of the most important curative factors. It is defined as a force of a symbiotic unity. On the other hand, the phenomenon of coherence is mainly understood as a harmonic connection of different members. Both phenomena, observed through sequences of the time in a group-analytic group, appeared as the result of the developmental processes of individuals in the group and of the group as a whole. The observation brought to a hypothesis that we are facing the parallelism of the individuation processes and group developmental process that have involved both – group-analytic and everyday interpersonal life. Group cohesion appeared as the result of the overcome chaotic paranoid position. The group, previously experienced as unknown and intimidating, becomes a symbiotic object, safe place gradually supporting the improving autonomy of individual members, mainly through important everyday relationships. Simultaneously and gradually group-analytic culture is changing from chaos and autism to equality, cohesion in the sense of “one mind”, toward coherence of interesting different members, from “forming and norming” to “performing”. It seems that group cohesion and group coherence are not phenomena *per se* but rather two positions in a group-developmental line, going from autism through cohesion and differentiation towards coherent interactive relationships.

Key words: group psychotherapy, group cohesion, coherence, group process

Od kaosa preko kozmosa do koinonije: Skupinskoanalitična razvojna linija

Peter Praper

Univerza v Ljubljani, Oddelek za psihologijo, Ljubljana

Povzetek: V raziskovanju zdravilnih dejavnikov v skupinski psihoterapiji se v zadnjih desetletjih največ pozornosti namenja fenomenu skupinske kohezije. Izraža se z občutki pripadnosti in navezanosti, ki se ponovno in ponovno potrjujejo kot pglavilni zdravilni dejavnik. Kohezijo lahko definiramo kot silo simbiotične enosti, ki zagotavlja varnost in bližino. Sodobne raziskave in analize pa vendar opozarjajo na možnost, da kohezija, če obvladuje skupino preveč in predolgo, konzervativno ohranja enoumje in je tako izvor obrambnega delovanja skupine. Zato nekateri koheziji ob bok postavljajo koherenco, ki jo prepoznavajo kot harmonično organizacijo sicer med seboj različnih posameznikov. Ko oboje – kohezijo in koherenco – opazujemo skozi skupinski terapevtski proces, skozi časovne sekvence, opazimo, da so ti premiki odvisni od zmožnosti za vzpostavljanje odnosov vključenih članov ter od narave in stopnje

*Naslov/Address: University of Ljubljana, Faculty of Arts, Department of Psychology, Aškerčeva 2, 1000 Ljubljana, Slovenia, e-mail: peter.praper@ff.uni-lj.si

njihove individualiziranosti. V skupini kot celoti se te karakteristike pokažejo v tipičnem razvojnem procesu. Na mestu je hipoteza, da gre za paralelizem individualizacijskih procesov posameznikov in skupinskih razvojnih procesov. Skupinska kohezija se razvije ob tem, ko vključeni presežejo začetno kaotično paranoidno pozicijo in si nudijo vzajemno razumevanje, varnost in oporo. Skupina, ki je sprva predstavljala neznan in zastrašujoč prostor, postane zatočišče simbiotične enotnosti. Preseganje kaosa in izkušnja pripadnosti prinašata medene tedne skupine, ki se jim je težko odpovedati na račun napredovanja posameznikov, tako v terapevtskih skupinah kot v vsakdanjem življenju. Vendar pa brez diferenciacije ni poti naprej. Kohezivna skupina se pogosto ob izgubljanju notranjih meja paranoidno zapira v polje odnosov z drugimi skupinami. V primerih, ko kohezivna skupina prične najprej tolerirati, nato pa tudi podpirati različnost pogledov, izkušenj, čustvenega doživljanja, prehaja iz kohezije v koherentno organizacijo. Pri tem ne gre le za toleranco, ampak predvsem za spoznanje, kako bogata je lahko izmenjava med različnimi ljudmi in nato tudi narodi, kulturami. Kaže, da kohezija in koherenca nista fenomena per se (kot je bilo oboje obravnavano do danes), ampak prej predstavljata dve poziciji v skupinski razvojni liniji, ki teče od avtizma, preko kohezije in nato po poteh diferenciacije do koherentnih interaktivnih odnosov.

Ključne besede: skupinska psihoterapija, skupinska kohezija, koherenca, skupinski procesi.

CC = 3313

During the 70s and 80s of former century the concept of cohesion in group therapy drew the attention of many researchers. It has been well known from psychoanalysis and other individual therapies, how the capacity for the deep and confidential relationship, on the side of the client and therapist, may be the crucial therapeutic factor. Speaking of group psychotherapy, this capacity is demanded by the majority of the members so that they can develop group cohesion. Truax (1961) asserted that cohesiveness may be the central factor of the analysis in the small group and successful groups are cohesive.

Yalom (1975) initiated a systematic research of the therapeutic factors. Followed by many other researchers, it was confirmed that group cohesion may practically be the prototypical therapeutic factor.

Yalom (1975) defined group cohesion as “the attractiveness of a group for its members” (Pines, 1998, p. 59). Larkin (1976, in Gazda, 1984) describes group cohesiveness as “the collective expression of personal belongingness [...] binding members emotionally to the common task as well as to one another, assuring greater stability of the group [...] and developing a shared frame of reference among group members that allows for more tolerance for diverse aims of group members. Obviously, then, a cohesive group is a stable and productive group that can be quite task or goal oriented” (p. 56). It seems to me that, in fact, Larkin is speaking of a coherent and not cohesive group or, at least, he seems to merge both of them. However, group cohesiveness may easily divert the behaving to conformity with the group norms that can produce an obstacle for reaching the goal of restoring or enhancing the member’s autonomy which, seen by Helen Durkin, is the primary goal of group psychotherapy (in Nicholas, 1984, p. 133). Durkin asserts that contact operations in

a group process should, at the same time, be a process of establishing boundaries. It seems to me that the problem, how to reach not ignorant but a tolerant and at the same time interactive relations with the respect for the boundaries of the other go far beyond the cohesion of the group.

Pines (1998) finally distinguished cohesion and coherence describing both as two different processes. They differ in the degree of organisation and differentiation. Cohesive group, although having a recognisable inner structure, is more like broken pieces glued or cemented together in one, while coherence indicates a specific organisation of different units into a spiritual, intellectual, aesthetic relation. It seems that only coherence is seen by Pines as the result of a developmental process. Pines is comparing cohesive sticking together with Bion's (1961) basic assumptions of performance of the group, functioning according to archaic emotions. Although creating a holding and containing environment, it does not develop where individuals merge. Cohesive group may, then, as well as protective, be a defensive one. On the contrary, when coherence is developed, the group went through "utopian stages, group illusion about enjoying in a sense of unity" (Pines, 1998, p. 65). A coherent group has replaced the emotional glue with a rational and functional self-organisation.

Pines (1998), as well as British developmental psychoanalysts in general, tries to avoid Mahler's concept of separation and individuation, although Anna Freud (1963), introducing the concept of the developmental lines, considered it as a prototypical developmental line, called *From dependency to emotional self-reliance and adult object relations*. Pines would rather see the processes of cohesion and coherence connected to attachment and identification phenomena. Quoting Lichtenberg (1982, in Pines, 1998) how psychoanalysis possibly over-states the separateness of man and his degree of independence, Pines continues to be critical about the western individualism and its reflection in psychoanalysis in the resistance to the interrelational concept. However, the question remains whether Mahler advocates an exaggerated individualism or rather explains both, the importance of a good enough symbiotic experience and the accomplished autonomy. Mahler's concept may well help to understand the difference among mature autonomy and immature and neurotic search for independence.

Ashbach and Schermer (1987) relate to Mahler's work, reminding us that Margaret Mahler herself was very careful not to generalise her findings prematurely to adults and groups. However, "the present authors hope that a tentative application of Mahler's point of view to group psychology does not constitute too great an inductive leap from a research paradigm intended primarily for the study of child development" (p. 92). According to Kauf (1977, in Ashbach & Schermer, 1987, p. 92), "aspects of symbiosis / separation-individuation processes are repeated in groups as both individual valences and group processes". Of course there are other aspects and developmental lines like attachment, mirroring, identification and last but not least cognitive development, mentioned by Pines (1998) as Robert Kegan's Neo-Piagetian approach, where we can find the link with ego-psychological concept

of hierarchically organised ego functions. Not that some of these theories may be incorrect, but it is probably true that the processes through different developmental lines are interrelated, influencing one another.

Clinical illustration

Masha and Darinka entered the group in September, after the summer leave. The group went into the fourth year of the existence. There were three old members included at the beginning and two that were included during the course of the group. Jan, a mail member, could not reach beyond his counter dependent attitude and prematurely ended the treatment after two years, in July.

Masha, a woman in her thirties, was a warm, over adapted person. Everybody liked her and she enjoyed the position where everyone needed her. Her mother, a highly professional woman, was, in her private life, constantly seeking Masha's emotional support, trying to use her daughter as a mothering object. Her father constantly demanded Masha's admiration. Although in fact he was an unsuccessful alcoholic, he was idealised by her. When her parents moved to the newly built house in the country, Masha was able to oppose the utopia that all of them could live happily there. So she remained in a flat in the city and took over the role of a guardian of the primary home and took care of her few years younger brother. Although she had a presentiment of co-dependency, she was not aware of her role of keeping the family together. She became more and more unhappy, having had some unsuccessful relations with mainly immature partners. She accepted an employment evidently below her capacities. Though she may be very responsible at her job, she has had a serious delay with her diploma at the faculty. Then she realized that she is avoiding growing up, and she decided to enter the therapy group.

Darinka has had a similar type of avoiding the role of an adult woman although her childhood was quite different. She had been raised by her grandmother and aunt at a small farm. Her mother left, went to the city to work in a factory and became a convinced proletarian. Coming to a visit during weekends, she kept developing more and more conflicts in relationships with her mother and sister. And yet there was an archetypical similarity – three generations, four women and no one created a lasting relationship with a man. Though Darinka as a child could not wait for her mother to come on Saturdays, her real primary object was her grandmother. Then the seven-years-old Darinka was simply moved to the town to live with her mother and started attending school. She reacted with severe separation crisis. Since then she has struggled with her psychopathology. In her late thirties she entered developmental analytical psychotherapy lasting nearly four years, and completed it quite successfully. She had realized how she was brought up by her grandmother in a feudal obedience and at the same time contaminated with her mother's proletarian ideology. Suffering because of the conflict among internalised objects and attitudes,

she created an idealised self representation of herself as an independent intellectual but still without her faculty diploma in her early 40s. However, she did not develop a capacity to create a successful partnership with a man or a cooperative attitude in a group. Both her private and professional life was affected by that deficit. Therefore, after a two-year break, she entered a therapeutic group, led by the same therapist.

Masha and Darinka entered a cohesive, defensive and relatively closed small (probably too small) group. Masha started with her over adapted and regressive behaviour, fitting well into a cohesive but somehow co-dependent culture of the group. Darinka, on the other hand, acted in a manner of a counter dependent behaviour, trying to discuss topics with the therapist only, offering herself to the group as a scapegoat. Before the group and her were ready to work out the problem, probably the most important therapeutic matter of the group, Darinka dropped out. Masha survived in a holding atmosphere and step-by-step started practising. With the support of the group she passed the last exam before the diploma at the faculty and after that somehow anxiously retired to a depressive dissatisfaction with her brother, parents and boyfriend. The group, ready to take sides, continued with giving the credit to the member, blaming those from outside. They would neither explore Masha's role nor the group cohesive but egalitarian position borderless inside and in a paranoid position towards outside.

Slowly the differentiation took place and some of the old members became preoccupied by their practicing and self-actualising activities. Although most of these processes took place outside the group it would slowly change the group culture, pushing in front the dynamic matrix. Masha forgot about the problems of her primary family and started working through the dynamics of another group of great importance to her. She was a founder member of a female singing group. Thirteen of them – more than a small group – have been performing for nearly ten years. As they were becoming more and more famous, the differentiation that took place was accompanied by increasing rivalry. A cohesive group before, they were confronted with a lack of boundaries and organisation so that a threat of a decay was on the door. Masha succeeded to connect the pieces together and accepted the challenge of how to learn to use the words “no, I, I want, I demand ...”. Parallel to the singing group it suddenly appeared within the therapeutic group as a training in action, in the family relations and in the interaction with her boyfriend. The unspoken contents became negotiable. Step by step Masha created the needed distance to her mother's expectations and her father's drinking problems without losing love for them. She decided to move to her boyfriend's town, taking the risk of living together, and left her brother on his own. At the same time her singing group was offered to give a gala concert for their tenth anniversary. She was the one introducing an acceptable musical director and promoting the idea how to create a very heterogeneous repertoire, exposing solo singers, duos, trios, larger parts of the group and the group as a whole. Masha described the outcome at the therapeutic group with enthusiasm. When the singing group discovered a harmony in differences, they experienced a

massive delight of coherence, and astonishing, they received unanimous support by their husbands or boyfriends.

That was the time when, in her third year of therapy, Masha started feeling ready to bring the group therapy to an end. The group also discovered the meaning of its defensive standstill in the pleasant mood of cohesion. They became not only tolerant but also eager to accept new members.

Conclusion

Clinical observation supports my idea about cohesion and coherence, not as two different processes but two stages of a same developmental spiral. I believe that we are speaking about the prototypical developmental line (see Figure 1) observed both in individual and group process, starting in autism or schizoid-paranoid position, going through dyadic merging or symbiotic type of dependency and border deficit towards the therapist, in a depressive position. As the communication spread around the group, the conflict fight for power in the group (what inevitably provokes feelings of envy, jealousy and rivalry) strengthen the feelings of being threatened and the members have to search for some alliance. In group therapy we can observe how the members make the transition from the therapist as an object to the group as a whole as an object. In this stage cohesion takes place. Primary, primitive emotional identification, the same feelings, one thought, group as a mother with the capacity to hold and contain – these are the signs indicating how the group represents to the majority of the members an object of the symbiotic, emotionally relevant expectations. And indeed, the cohesive group offers a lot of understanding and grandiose mirroring. It represents the basic matrix offering the atmosphere of basic trust, the feeling of belongingness as the correlate of the attachment. However, the group during this honeymoon stage may also be egalitarian, with border deficit within the group, provoking the fear of engulfment. It is naturally preceded by counter dependent struggle for self. The fascination may be followed by the disappointment and in this group position we can face the danger of the destruction of the group, the phenomenon of late dropouts or premature counter dependent conclusion of some member, searching the way through the distorted, self-sufficient pseudo individuality in the sense of independence instead of autonomy. Group therapeutic groups are composed of mainly disturbed persons, often avoiding the responsibility as an aspect of autonomy, and that is why prolonged dependency or the forms of pseudo individuation are seen more often.

The group that succeeds in developing beyond egalitarian and counter dependent position not only tolerates but supports and engages specific differences among members. Some research work of Vec (2001), a student of mine, showed how this process may be far from linear succession of changes, however, through the chaos of the events usually an order in the sense of a spiral movement in the opened auto

poetic system emerges. Usually the sense of belonging that goes beyond a cohesive group as the attachment goes beyond symbiosis, gives the individual the strength to take over the responsibility of an autonomous person, responsible also for the group as a whole. As we know from developmental psychology, moving to the advanced developmental phase does not mean that the characteristics of the previous phase have been extinguished. Although not dominant any more, the previous system of actions remains as a possibility. Therefore in a coherent group the experience of cohesion not only remains present but continuously supports connection and belongingness while the differentiation in a state of coherence decreases the fear of engulfment and loss of identity. As for the individual, the same for the group as a whole this developmental level brought the capacity of regression in the service of the ego/group organisation. The experience of cohesion will prevent the group members to move to egocentric individualism and the group to ethnocentrism (in relation to other groups), as well as the experience of coherence will prevent dedifferentiation in the moments of surrendering to cohesive unity. The paradox between connectedness and differentiation is resolved by the integration.

In 1994 at a congress in Madrid, I entered a discussion with Malcolm Pines, when he first claimed for coherent instead of cohesive relationships in the macro system of international relations. My idea how those who were living in egalitarian systems need some time for the process of individuation to become potent to establish a harmony in coherence was hardly understood. Then I offered the metaphor how a child learns to sing in a group with more voices singing. First he would merge, imitating other voices, then, gradually he learns his specific voice and after that becomes able to fit in the harmony. Masha has, parallel to the group therapy curative environment, reached her maturity literally through these processes and helped the group to move forward.

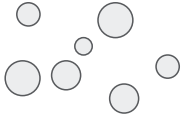
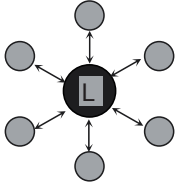
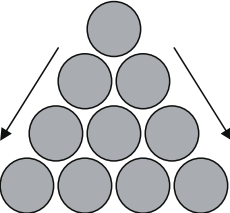
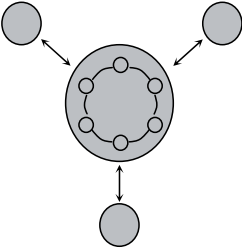
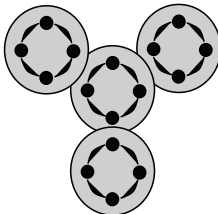
Negativism		
Forming		<p>DEPENDENCE (individual in the relation to the leader) “Being” principle</p>
<p>Storming Fight – Flight</p>		<p>INTRAGROUP CONFLICT COUNTER- DEPENDENCE (individuals) “Doing” principle action – reaction egocentrism</p>
Norming		<p>CODEPENDENCE ON GROUP AS A WHOLE inner fusion and paranoid reaction to outside: ethnocentrism COHESION</p>
Performing		<p>INTERACTION INTERDEPENDENCE in the sense of mutual solidarity of different individuals (inside) and groups as whole (outside) COHERENCE</p>

Figure 1. Group developmental process. Remarks: (i) The advanced developmental achievement does not cause the extinction of the previous phase. (ii) The quality of cohesion remains important in a coherent group as well as the sense of belonging, however, there is a new sense of individual responsibility. (iii) The experience of coherence prevents the fear of engulfment as well as the danger of individualistic alienation. (iv) The experience of coherence enables the emphatic interaction without losing the differentiation.

References

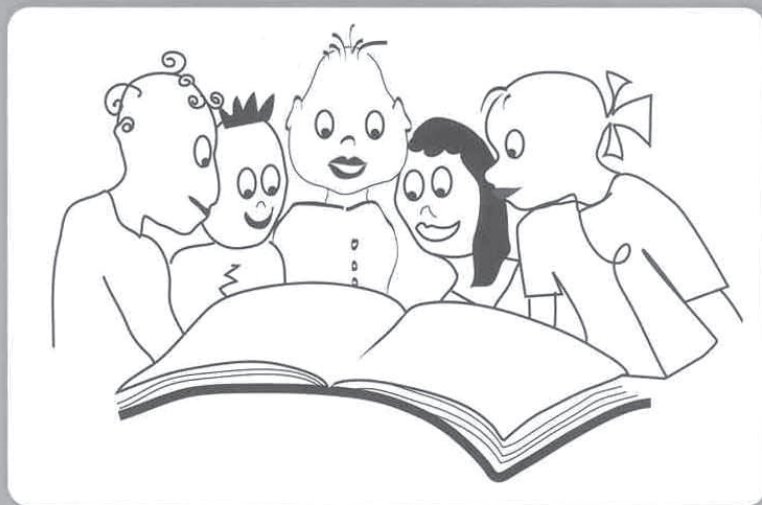
- Ashbach, C., & Schermer, V. L. (1987). *Object relations, The Self and the Group*. New York: Routledge.
- Bion, W. R. (1961). *Experiences in Groups*. London: Tavistock.
- Freud, A. (1963). The concept of developmental lines. *Psychoanalytic Study of the Child*, 18, 245–265.
- Gazda, G. M. (1984). *Group counselling: A developmental approach*. Boston: Ally and Bacon.
- Nicholas, M. W. (1984). *Change in the context of group therapy*. New York: Bruner-Mazel Publishers.
- Pines, M. (1998). *Circular reflections: Selected papers on group analysis and psychoanalysis*. London: Jessica Kingsley Publishers.
- Truax, C. B. (1961). The process of group psychotherapy. *Psychological Monograph*, 75(511).
- Vec, T. (2001). *Nekateri vidiki supervizijskega procesa pedagoških delavcev [Some aspects of supervision process with educational workers]*. Neobjavljeno specialistično delo, Univerza v Ljubljani, Filozofska fakulteta [Unpublished manuscript, University of Ljubljana, Faculty of Arts].
- Yalom I. D. (1975). *The theory and practice of group psychotherapy*. New York: Basic Books.



Center za pedagoško izobraževanje Filozofske fakultete

**MELITA PUKLEK LEVPUŠČEK
BARICA MARENTIČ POŽARNIK**

SKUPINSKO DELO ZA AKTIVEN ŠTUDIJ



Ljubljana, 2005

Povezava med Eysenckovimi osebnostnimi potezami in spominsko obnovo

*Igor Areh**

Univerza v Mariboru, Fakulteta za varnostne vede, Ljubljana

Povzetek: Namen raziskave je bil ugotoviti stopnjo povezanosti med osebnostnimi potezami in količino ter točnostjo spominskega priklica, kar bi omogočilo oblikovanje modela za napovedovanje verodostojnosti pričanja očitvev kaznivih dejanj. Raziskovalne hipoteze temeljijo na dosedanjih raziskovalnih in teoretičnih pričakovanjih: visok nevroticizem, psihoticizem in visoka introvertnost so povezani z manjšo točnostjo in količino spominske obnove. V raziskavi je sodelovalo 280 študentov, starih od 18 do 21 let. Po uporabi osebnostnega vprašalnika EPQ-R so si udeleženci ogledali posnetek telesnega napada na žensko, ki se je končal z ropom. Po sedmih dneh sem s čeklisto ugotavljal točnost in količino spominske obnove. Ugotovil sem, da psihoticizem ni povezan s točnostjo in količino spominske obnove. Ekstravertnost pri moških ne korelira z merjenimi značilnostmi spominske obnove, pri ženskah pa je pozitivno povezana le s količino spominske obnove. Visok nevroticizem je pri ženskah povezan z manj točno in količinsko siromašnejšo obnovo. Pri moških je visok nevroticizem povezan le z manjšo količino obnovljenih podatkov. Kaže, da lahko pri ekstravertnih ženskah pričakujemo večjo količino spontano navedenih informacij o dogodku, pri tem ni upravičeno pričakovati bolj točne spominske obnove. Dobljeni korelacijski koeficienti so nizki in sklepam, da uporabnega modela za napovedovanje verodostojnosti pričanja ni mogoče izoblikovati. Šibkost raziskave je manjša reprezentativnost vzorca, saj je povprečna vrednost osebnostne poteze ekstravertnost nižja od normativnih vrednosti. Ugotovitve raziskave so pomembne za učinkovitejšo presojanje verodostojnosti pričanja očitvev, prispevajo pa tudi k razjasnitvi odnosov med osebnostnimi potezami in spominsko obnovo. Prednost raziskave je v višji ekološki veljavnosti, saj je izvedena na način, ki posnema dogajanje v resničnosti.

Ključne besede: osebnostne lastnosti, Eysenckov vprašalnik osebnosti (EPQ-R), spominska obnova, pričanje očitvev

Connection between Eysenck's personality traits and memory recall

Igor Areh

University of Maribor, Faculty of Criminal Justice and Security, Ljubljana, Slovenia

Abstract: The research examines the efficiency of episodic memory in the context of eyewitness memory. The purpose of the research was to ascertain the connection between personality traits and memory recall, what may allow us to form a model of reliability prediction of eyewitness testimony. It was hypothesized that high neuroticism, psychoticism and high introversion are connected with low accuracy and quantity of memory recall. Participants were undergraduate students ($N = 280$) between

* Naslov/Address: Fakulteta za varnostne vede, Univerza v Mariboru, Kotnikova 8, 1000 Ljubljana, e-mail: igor. areh@fvv.uni-mb.si

the ages of 18 and 21. Personality traits were measured by Eysencks' Personality Questionnaire (EPQ). After the application of questionnaire a video about a robbery assault on a female was shown. Seven days later the accuracy and quantity of memory recall were tested with an extensive checklist. It was established that extraversion is not correlated with memory recall in men, while in females a positive correlation was found. Among females, high neuroticism was correlated negatively with the accuracy and quantity of memory recall, and in men only a negative correlation with quantity of memory recall was found. All correlation coefficients were weak, therefore it was impossible to form an applicable model of reliability prediction of eyewitness testimony. The main weakness of the research presented was the sample anomaly, which was evident in noticeable deviation in the average level of extraversion when it was compared with the expected standard for the Slovenian population.

Key words: personality traits, Eysenck Personality Questionnaire (EPQ), memory recall, eyewitness testimony

CC = 4230, 3120

Raziskovanje povezav med osebnostnimi dejavniki in učinkovitostjo spominske obnove sega v obdobje 19. stoletja. Sprva je bila pozornost raziskovalcev usmerjena na preučevanje vpliva, ki ga imajo psihične motnje ali bolezni na spominski priklic. Tu velja omeniti Ribojovo preučevanje amnezij (Ribot, 1882, po Schacter, 1996) in Freudovo razlaganje izrinjanja zavestnih vsebin v nezavedno (Freud, 1909). Kasneje so začeli preučevati tudi vplive lastnosti »normalne« duševnosti na priklic informacij iz dolgoročnega spomina, vendar je le redke raziskovalce zanimal vpliv osebnostih potez ali dimenzij na spominsko obnovo. Raje so raziskovali spominske zmožnosti, kot bi bile neodvisne od osebnosti. V tej skupini najdemo veliko vplivnih imen: Ebbinghaus, Binet, Henri, Semon, Bartlett itn. V 60. letih prejšnjega stoletja je prišlo do prelomnice v raziskovanju. Število raziskav se je naglo povečalo, vzrok za to pa je bila novo nastala računalniška metafora. Njena era se je začela leta 1958 z objavo članka o računalniškem programu za igranje šaha, ki so ga napisali Allen Newell, Cliff Shaw in Herbert Simon (Hergenhahn, 2001). V tem času se je zgodilo še nekaj pomembnega – če je bila veljavnost psiholoških spoznanj do 60. let v veliki meri omejena s sterilnostjo laboratorijskega okolja, je tedaj, v želji po doseganju višje ekološke veljavnosti, psihologija stopila na pot k znanosti, ki je uporabna v vsakdanjem življenju. Razmahnilo se je preučevanje t. i. vsakdanjega spomina (spomina na vsakdanje dogodke), kar so v veliki meri spodbudila številna življenjska, praktična vprašanja, ki so se pojavljala pri delu sodišč in preiskovalcev kaznivih dejanj. To so bila vprašanja kot npr., kako oceniti verodostojnost pričanja, ali je izkazovanje večje gotovosti v spomin povezano tudi z večjo točnostjo obnove, katere podrobnosti v spominski obnovi so bolj izpostavljene popačenju in pozabi in katere manj, kaj vpliva pri opisovanju resničnih dogodkov na kakovost in količino zapomnjenih informacij, ali osebe, ki so vpletene v dogodke kot žrtve, bolj verodostojno opisujejo dogajanje ali pa so boljši tisti, ki niso vpletene v dogajanje, itd. Treba je opozoriti, da se zgodovina raziskovanja vsakdanjega spomina ni začela šele v šestdesetih letih

prejšnjega stoletja. Za pionirje tega področja štejemo McKeena Cattela, ki je leta 1893 izvedel poskus, v katerem je študentom zastavil vprašanja, ki jih lahko pričakujemo na sodni razpravi, in ugotovil, da je spomin na vsakdanje, nepomembne dogodke zelo varljiv (Bartol in Bartol, 2006). Alfred Binet je leta 1900 izdal knjigo »La Suggestibilit e« (Wells, Memon in Penrod, 2006), Louis William Stern je leta 1904 za el z zaporedjem zanimivih poskusov s področja pri anja o ividcev (Wells idr., prav tam), Hugo M nsterberg (1908) je istega izdal knjigo »On the witness stand« itn. Omenjeni raziskovalci so bili nekako nerazumljeni med sodobniki, saj se strokovna javnost ni zavedala, da je pri anje o ividcev pogosto nezanesljivo, tako je bil tudi pomen njihovega dela za delo sodi  v veliki meri prezrt.

V  lanku se osredoto am na eno izmed danes zapostavljenih podro ij raziskovanja vsakdanjega spomina – na ugotavljanje stopnje povezanosti med to nostjo (verodostojnostjo) in koli ino spominskega priklica na eni ter osebnostnimi potezami na drugi strani. Omenil sem, da se je v drugi polovici 20. stoletja obseg raziskav u inkovitosti spomina bistveno raz iril in poglobil. Vendar pa to ne velja za raziskovanje povezave med osebnostnimi potezami (izven okvira psihopatologije) in spominskim priklicem. Teh raziskav je malo, zato so omenjene povezave slabo raziskane, obstoje a spoznanja so nejasna ali pa si nasprotujejo. Kljub temu lahko na splo no ugotovimo, da se v raziskavah, kjer so potrdili obstoj povezanosti med osebnostnimi potezami in dolo enimi zna ilnostmi spominske obnove, vztrajno pojavljajo  ibki korelacijski koeficienti (Wells in Olson, 2003). Poglejmo dosedanje ugotovitve.

Introvertiranost in ekstravertiranost

Ward in Loftus (1985) sta sku ala raziskati odnos med dimenzijo ekstravertnost – introvertnost in to nostjo ter koli ino spominskega priklica. Pri tem sta uporabila Myers-Briggsov vpra alnik osebnostnih tipov, ki izhaja iz prilagojene Jungove teorije psiholo kih tipov (Myers-Briggs in Myers, 1980). Na osnovi tedanjih raziskovalnih rezultatov sta predpostavljala, da je introvertnost povezana s slab o spominsko obnovo dogodkov. Pred njima je Eysenck (1981) ugotovil, da imajo introverti vi je osnovno vzburjenje osrednjega  iv evja, in po njegovem mnenju je visoko osnovno vzburjenje, predvsem pri zaznavanju hitro odvijajo ih se dogodkov, ovira za u inkovito vkodiranje informacij v dolgoro ni spomin. Nasprotno velja za ekstravertne osebe, saj imajo ni je osnovno vzburjenje, zato naj vkodiranje podatkov ne bi bilo moteno. Ker so ekstraverti usmerjeni v intenzivnej e do ivljanje svoje okolice, so bolj pozorni na dogajanje v okolju in zato bolj »uporabni« v vlogi o ividcev (Ward in Loftus, prav tam). Domnevo o usmerjenosti ekstravertov v intenzivnej e do ivljanje okolja potrjujejo tudi visoke korelacije med ekstravertnostjo in dvema Zuckermanovima dimenzijama »iskanje novosti« ter »odprtost za izku nje« (Aluja, Garcia in Garcia, 2003).

Ward in Loftus (1985) sta se oprla na Eysenckove domneve o negativni povezavi med introvertnostjo in točnostjo spominske obnove (Eysenck, 1981) in dejansko ugotovila obstoj negativne povezave med introvertnostjo in točnostjo pričanja, vendar je bil korelacijski koeficient nizek in statistično ni bil pomemben. Menila sta, da je nižja točnost pričanja pri introvertih posledica njihove večje dovzetnosti za vpliv podogodkovnih informacij (gre za informacije, do katerih pridemo po dogodku, te se pogosto spojijo z izvornimi spomini in jih popačijo). Večja dovzetnost ali sugestibilnost za podogodkovne informacije naj bi bila posledica višjega osnovnega vzburjenja introvertov, ki moti proces vkodiranja informacij (Ward in Loftus, 1985). Zaradi motenj prihaja do nastajanja spominskih vrzeli, ki pa niso trajne ali stabilne. Ravno nasprotno – ob vsakem priklicu podatkov v zavest skušamo vrzeli zapolniti z neko smiselno vsebino (konfabulacije). Eden najbolj prikladnih virov za zapolnjevanje informacijskih vrzeli so podogodkovne informacije. Zaradi nezavedne težnje po zapolnjevanju spominskih vrzeli se pojavlja večja sugestibilnost, ta pa dodatno znižuje verodostojnost spominske obnove (Ward in Loftus, prav tam). E. Loftus (1996) je kasneje dodala domnevo, da so introverti bolj sugestibilni kot ekstraverti tudi zaradi nižje samopodobe, ki vodi v občutek manjše gotovosti v lastni spomin. Negotovost pa zlahka povežemo z večjo dojemljivostjo za mnenja drugih oz. z večjo dojemljivostjo za podogodkovne informacije.

Na Eysenckovo teorijo o vlogi ravni vzburjenja osrednjega živčnega sistema se je oprl tudi Schwartz (1975), ki je menil, da osnovno vzburjenje vpliva na organiziranost spominske mreže in tako posredno tudi na dostopnost podatkov v njej. V dveh raziskavah je potrdil hipotezo, da se z zvišanjem vzburjenja zniža raven procesiranja podatkov iz semantične na fizično oz. na raven zaznavanja dražljajev. Zato se osebe z višjim osnovnim vzburjenjem (introvertnost ali visok nevroticizem) bolj osredotočajo na fizikalne lastnosti dražljajev, osebe z nižjim vzburjenjem (ekstravertnost ali nizek nevroticizem) pa na pomen zaznav. Na osnovi tega je Schwartz (prav tam) sklepal, da si introverti bolj zapomnijo podrobnosti dogajanja, vendar jih ne osmislijo tako dobro kot ekstraverti. Podrobnosti, ki niso smiselno povezane v širšo celoto, so bolj izpostavljene procesom pozabljanja, zato je pričakovati, da so introverti manj učinkoviti pri spominski obnovi kot ekstraverti. Schwartzove rezultate so pozneje potrdili tudi drugi raziskovalci (Libkuman, Griffith, Wines, Dickel in Doty, 1998).

Rezultati raziskav s konca prejšnjega stoletja opozarjajo tudi na zelo verjeten obstoj pozitivne korelacije med ekstravertnostjo in nagnjenostjo k procesiranju pozitivnih čustev (Gomez, Gomez in Cooper, 2002; Rustig in Larsen, 1997), kar je predvidel tudi Hans J. Eysenck (1970). Zato je za ekstravertne osebe značilna tudi nagnjenost k priklicu večjega deleža prijetnih spominov in dajanju pozitivnih mnenj. Ta nagnjenost ne prispeva k večji točnosti spominske obnove, saj vodi v pristransko obdelavo, vkodiranje in tudi v pristranski priklic informacij iz dolgoročnega spomina (Rustig, 1999).

Večina raziskav o povezavi med ekstravertnostjo ali introvertnostjo na eni strani in učinkovitostjo spomina na drugi je bila narejena z uporabo asociativnih

parov, zato je posploševanje ali prenos ugotovitev na področje pričanja očitvecev vprašljiv. Pri raziskovanju z asociativnimi pari si udeleženci npr. ogledajo seznam besednih parov, nato jih po določenem času priključijo v zavest. Pri tem lahko ima ena od besed v paru vlogo ključne ali sprožilne besede, ki omogoči priklic druge besede. Gre torej bolj za raziskovanje zmožnosti prepoznavanja dražljajev in ne za raziskovanje priklica življenjskih dogodkov. Vendar pa ugotovitve »asociacijskih« raziskav niso v neskladju z ugotovitvami raziskav verodostojnosti pričanja. Rezultati obeh vrst raziskav so si namreč podobni, kar je razvidno tako pri analizi spominskega priklica dogodkov kot tudi pri analizi spomina na izmišljene zgodbe (Lieberman, 2000).

Nevroticizem in anksioznost

O učinkovitosti vkodiranja informacij v spominsko mrežo dolgoročnega spomina odločata tudi raven čustvenega vzburjenja, ki se ob zaznavi dražljaja inducira v osrednjem živčevju, in čustvena valenca dražljaja. Informacije, ki imajo intenziven čustveni naboj, se bolj uspešno vkodirajo v spomin kot informacije z neizrazitim in nevtralnimi čustvenim nabojem (Leigland, Schulz in Janowsky, 2004).

Čeprav se je raziskovanje vpliva nevroticizma na spominski priklic začelo pred več kot 30 leti, je zveza še vedno nejasna. Kljub temu se počasi in vedno bolj uveljavlja spoznanje, da je visok nevroticizem povezan s pogostejšim priklicem podatkov, ki so povezani z negativnimi čustvi ali izkušnjami (Mansell, 2004; Ruiz-Caballero in Bermudez, 1995), in s precenjevanjem jakosti negativnih čustev, ki smo jih doživeli v preteklosti (Levine in Safer, 2002). Vse bolj se potrjuje hipoteza o učinku osebnostno skladnega priklica (angl. *personality-congruent recall effect*), ki predpostavlja, da je spoznavno procesiranje informacij z negativnim pomenom pospešeno pri osebah, ki dosegajo visoke vrednosti na Eysenckovi lestvici nevroticizma (Martin, Ward in Clark, 1983). Če povzamem, osebe z visokim nevroticizmom se pogosteje kot ekstravertne osebe spominjajo negativnih izkušenj in pri tem pogosteje dramtizirajo lastna doživetja. Zaradi pogostejšega obnavljanja spominov bi pričakovali bolj točne in podrobne spomine, vendar se to praviloma ne dogaja, saj so korelacijski koeficienti med nevroticizmom in količino ali točnostjo spominske obnove običajno šibki in negativni (Hermans idr., 2004). Morda se pri introvertnih osebah spomin na slabe dogodke, zaradi vpliva doživljanja intenzivnih negativnih čustev, hitro pretvarja v t. i. spomin bliskavice (angl. *flashbulb memory*). Za slednje je značilno, da je oseba zelo prepričana v točnost lastnega spomina, ki je poln podrobnosti, kljub temu pa je poln popačenj in precej odstopa od resničnosti.

Kaže, da tudi anksioznost vpliva na priklic podatkov iz epizodičnega spomina. Shapiro in Penrod (1986) sta ugotovila, da obstaja povezava med kronično anksioznostjo kot osebnostno potezo in točnostjo prepoznave osumljencev v prepoznavni vrsti. Visoka anksioznost je povezana z manjšim številom zmot pri prepoznavanju osumljencev kaznivih dejanj. Nasprotno velja za osebe z nizko anksioznostjo, ki jo

povezujeta z večjim številom zmotnih prepoznav. Avtorja sklemeta, da se manj anksiozne (brezskrbne) osebe manj obremenjujejo z morebitnimi negativnimi posledicami zmotnih izjav, da se odzivajo manj premišljeno in verjetno se manj anksiozne osebe čutijo tudi manj odgovorne za sodelovanje s preiskovalci kaznivih dejanj. Nadalje sklepata, da učinkovitost prepoznavanja osumljencev oz. spominske obnove z višanjem ravni vznburjenja ali anksioznosti najprej narašča. Ko raven anksioznosti doseže mejno vrednost (plato), je učinkovitost delovanja spomina optimalna. Nadaljnje višanje anksioznosti pa učinkovitost spomina le še znižuje. Ta odnos lahko najbolj preprosto ponazorimo z Yerkes–Dodsonovim zakonom oz. z obrnjeno U krivuljo, s katero sta Yerkes in Dodson (1908) opisala učenje »plešočih« miši, kasneje pa so z njo začeli ponazarjati odnos med učinkovitostjo spoznavnih sposobnosti in ravno vznburjenja v osrednjem živčevju.

Anksioznost vpliva tudi na usmerjanje pozornosti. Anksiozne osebe usmerjajo pozornost predvsem v tiste podrobnosti dogajanja, ki se jim zdijo ogrožajoče (Mogg in Bradley, 1998; Narby, Cutler in Penrod, 1996). Kasneje, pri spominski obnovi, pa se ozaveščanju ogrožajočih podrobnosti izogibajo (Rohner, 2002). Pri anksioznih osebah se zaradi izrazitejših občutkov ogroženosti pogosto pojavi učinek cevnege ali predorskega pomnjenja (t. i. osredotočenje na orožje) in oseba si zapomni predvsem tisto, kar je zaznala kot ogrožajoče, npr. nož, palico, pištolo, masko ipd. Drugih podrobnosti, ki je niso ogrožale, se slabo ali pa sploh ne spominja. Visoko anksiozne osebe se v praksi izkažejo kot manj uporabne priče tudi zaradi bojzani pred morebitnimi negativnimi posledicami pričanja. Za anksiozne posameznike je namreč ugotovljeno, da precenjujejo verjetnost, da bi se jim zgodilo nekaj slabega (Kverno, 2000). Poleg tega so nagnjeni k temu, da si nejasne, dvomljive dražljaje ali okoliščine pojasnijo kot bolj negativne ali slabše, kot so v resnici (Brendle in Wenzel, 2004; Stopa in Clark, 2000; Wenzel, Pinna in Rubin, 2004). Nekateri tudi ugotavljajo, da imajo anksiozne osebe težave pri priklicu informacij, ki so povezane z intenzivnimi čustvi, saj je njihov spominski priklic manj točen in bolj skromen (Terry in Burns, 2001). Podobno velja za depresivna stanja, v katerih se znajdejo žrtve kaznivih dejanj ali nesreč. Depresivnost pospeši priklic informacij, ki imajo negativen pomen (Reidy, 2004; Reidy in Richards, 1997), kar lahko opazimo tudi v spremembi prepoznavanja čustvenih izrazov na obrazih. Depresivni in anksiozni posamezniki so nagnjeni k temu, da resne ali nevtralne obrazne izraze zaznajo kot jezne ali nenaklonjene (Gilboa-Schechtman, Erhard-Weiss in Jeczemien, 2002). Domnevamo lahko, da anksiozna (nevrotična) oseba, ki je nagnjena k zaznavanju drugih oseb kot nenaklonjenih ali celo sovražnih, vzpostavlja manj konstruktivne odnose. Anksiozna priča ali žrtev kaznivega dejanja bo verjetno zato vzpostavila slabši odnos s preiskovalci, ki jo intervjuvajo. V takih primerih so lahko podatki, ki jih žrtev posreduje, skopi, lažni ali zavajajoči.

Slabši spominski priklic nevrotičnih oseb ni nujno povezan z anksioznostjo. Nekateri raziskovalci dokazujejo, da se anksioznost nevrotičnih oseb poviša zaradi

občutkov ogroženosti, ki se pojavijo med raziskavo, ko uporabljamo preizkušnje ali teste (Sarason, 1972, po Dobson in Markham, 1993). Višja anksioznost povzroči, da se tok pozornosti razdeli na vsiljive misli ali bojazni, ki se pojavljajo med preizkušnjo, in na dejavnosti, povezane z opravljanjem preizkušnje. Zaradi delitve pozornosti je vkodiranje v dolgoročni spomin manj učinkovito, s tem pa je slabši tudi spominski priklic. Žal tudi tu ne gre brez nasprotij. Tako sta Dobson in Markham (1993) ugotovila, da se pri anksioznih osebah kakovost spominskega priklica izboljša, če vedo, da so na preizkušnji. Kljub temu prevladuje mnenje, da anksioznost praviloma slabi učinkovitost spominskih funkcij.

V Sloveniji raziskav, ki bi ugotovljale odnose med posameznimi značilnostmi spominske obnove in osebnostnimi potezami, nisem zasledil. Razumevanje teh odnosov je pomembno, saj bi omogočilo bolj zanesljivo ocenjevanje verodostojnosti pričanja žrtev in očitidcev v predkazenskih in kazenskih postopkih ter tudi v civilnih postopkih sodišč. Zato sem zasnoval in izpeljal raziskavo, pri kateri je bila osnovna ideja izdelava modela za napovedovanje verodostojnosti pričanja. Predhodno raziskavo s tega področja sem zaključil z domnevo, da modela, ki bi bil uporaben za delo sodišč, kar pomeni, da bi imel visoko napovedno vrednost, ni mogoče narediti (Areh, 2003). Pri ponovljeni raziskavi, v kateri sem spremenil metodologijo vrednotenja spominske obnove, sem ugotovil podobno nizke korelacijske koeficiente in potrdil domnevo o neobstoju uporabnega modela (Areh, 2007). Pri zadnji raziskavi sem postavil tri hipoteze in sicer:

- Ekstravertne osebe imajo bolj točen in količinsko bolj obsežen spominski priklic kot introvertne osebe.
- Osebe z visokim nevroticizmom prikažejo manj točno in količinsko skromnejšo spominsko obnovo kot osebe z nižjimi vrednostmi nevroticizma.
- Osebe z visokim psihoticizmom imajo manj točno in količinsko slabšo spominsko obnovo dogodka.

Pri prvih dveh domnevah sem se naslonil na rezultate drugih raziskovalcev. Pri domnevi o vplivu psihoticizma pa se nisem mogel opreti na ugotovitve predhodnih raziskav, ker jih preprosto nisem zasledil (naj spomnim, da v tem prispevku ne obravnavam klinične populacije). Zato sem domnevo postavil na temelju naslednjih teoretičnih pričakovanj: ker je visok psihoticizem povezan s samotarstvom, z neobčutljivostjo za potrebe drugih, s sovražnostjo in agresivnostjo tudi do bližnjih oseb, z odsotnostjo občutkov krivde ipd. (Eysenck in Eysenck, 1997), je težko pričakovati, da bo takšna oseba motivirana za sodelovanje s preiskovalci kaznivih dejanj oz. z raziskovalcem. Osebe z visokim psihoticizmom naj bi zato imele manj točno in bolj skopo spominsko obnovo dogodka.

Metoda

Udeleženci

V raziskavi je sodelovalo 280 rednih študentov prvega letnika Fakultete za varnostne vede, udeležba je bila prostovoljna. V vzorcu je bilo 161 ali 57 % žensk in 119 oz. 43 % moških. Velika večina (231 ali 83 %) udeležencev je bilo starih od 18 do 20 let. Prvi letnik študentov je bil izbran zato, ker je njihovo znanje o preiskovanju kaznivih dejanj majhno, o problematiki pričanja očitvidcev pa ne vedo ničesar.

Spremenljivke

Kot napovedne spremenljivke, s katerimi sem skušal napovedati točnost in količino spominske obnove, sem uporabil osebnostne dimenzije psihoticizem, ekstravertnost in nevroticizem.

Odvisni spremenljivki sta bili dve:

1. *Točnost obnove (TO)*, ki kaže stopnjo ujemanja med priklicanimi in resničnimi podrobnostmi. Izračunana je z enačbo:

$$TO = \frac{\sum X_{rpp}}{\sum X_p + \sum X_{npp}}$$

TO ... točnost spominske obnove.

$\sum X_{rpp}$... vsota resničnih podrobnosti posameznika.

$\sum X_{np}$... vsota vseh resničnih podrobnosti v gradivu (konstanta).

$\sum X_p$... vsota neresničnih podrobnosti posameznika (napake).

Točnost spominske obnove je prikazana kot razmerje med vsoto resničnih podrobnosti, ki jih oseba prikliče iz spomina, in seštevkom dveh delnih vsot – vsote vseh resničnih podrobnosti, med katerimi oseba izbira na seznamu (gre za konstanto s številčno vrednostjo 85), in vsote neresničnih podrobnosti, ki jih je oseba označila na seznamu. Številčna vrednost točnosti obnove se giblje med 0 in 1. Če vrednost pomnožimo s 100, je točnost priklica izražena v odstotkih.

2. Količina spominske obnove (*KO*):

$$KO = \frac{\sum X_{rpp} + \sum X_{npp}}{\sum X_p}$$

KO ... količina obnove.

ΣX_{rpp} ... vsota resničnih podrobnosti posameznika.

ΣX_{npp} ... vsota neresničnih podrobnosti posameznika.

ΣX_p ... vsota vseh podrobnosti, navedenih v seznamu.

Količino obnove sem opredelil kot razmerje med vsoto vseh resničnih in neresničnih podrobnosti ter vsoto vseh podrobnosti (resničnih in neresničnih), med katerimi oseba izbira na seznamu za preverjanje spominske obnove (konstanta z vrednostjo 101).

Inštrumenti in gradivo

Za merjenje osebnostnih potez sem uporabil Eysenckov vprašalnik EPQ-R (Eysenck in Eysenck, 1997). Točnost in količino spominske obnove sem ugotavljal s posebnim ocenjevalnim seznamom (»čeklisto«), s pomočjo katerega sem preverjal prisotnost podrobnosti v posameznikovi spominski obnovi. V seznam so bile vključene podrobnosti, ki so jasno razvidne in ključnega pomena za razumevanje dogodka.

V raziskavi je bil uporabljen dvominutni posnetek zaigranega prizora, ki prikazuje roparski napad. Posnetek je bil narejen s pomočjo mlajših poklicnih igralcev, ki sta širši javnosti neznana. Gradivo je bilo posneto na Fakulteti za varnostne vede, v prostoru, ki je študentom nedostopen in neznan. Za snemanje sem uporabil kamero, nameščeno pod stropom, tako kot so nameščene nadzorne kamere. Ker je bil posnetek udeležencem predstavljen kot posnetek varnostne kamere, je sestavljen iz enega kadra, posnetega v slabših svetlobnih pogojih, slika je barvna, dodan je zvočni zapis. Dogajanje poteka na stopnišču, začne se z naključnim srečanjem moškega in ženske. Moški se sprva vede nenasilno in prosi za manjšo vsoto denarja, ki pa je ženska ni pripravljena dati. Prošnje se počasi stopnjujejo v zahteve, moški postaja vedno bolj nesramen in dogajanje doseže vrhunec v telesnem napadu na žensko. Ropar žensko dvakrat udari v predel trebuha, pri tem se ženska sesede na tla, moški ji strga torbico z ramen in pobegne iz kadra.

Posnetek je bil narejen in pred ogledom predstavljen tako, da pri gledalcih vzbudi prepričanje v resničnost dogodka. S tem sem skušal doživljanje gledalcev čim bolj približati doživljanju pri opazovanju resničnega dogodka (induciranje čustvenega vznemirjenja, doživljanja dramatičnosti in resnosti dogajanja, vživljanje v žrtev ipd.). Ugotovljeno je namreč, da so spomini na resnične dogodke, ki pričo vsaj zmerno čustveno aktivirajo, bolj trajni in točnejši (Niedźwieńska, 2003).

Seznam za preverjanje spominske obnove je bil razdeljen na štiri dele. V prvem delu je bilo navedenih 20 podrobnosti posnetega dogajanja. Za točen opis vsake od podrobnosti je bilo naštetih več odgovorov (vsaj pet), med katerimi so morali udeleženci izbrati in označiti pravega. Če se podrobnosti ali odgovora niso spomnili

ali pa so bili negotovi, so označili odgovor »ne vem«. Drugi del seznama preverjanja spominskega priklica je bil namenjen opisu moškega, ki je napadel žensko. Udeleženci so videz osebe opisali z 28 podrobnostmi, pri vsaki od njih je bilo navedenih več možnih odgovorov, pravilen je bil le eden. V tretjem delu so udeleženci opisovali videz žrtve oz. ženske. Navedenih je bilo 29 podrobnosti opisa ženske, udeleženci so jih tudi v tem primeru določili z izbiro enega med več možnimi odgovori. Na koncu seznama za ugotavljanje kakovosti in količine spominske obnove so bile dodane 7-stopenjske Likertove lestvice, s katerimi so udeleženci ocenili značilnosti spomina na dogodek (npr. spomin na opazovani dogodek lahko ocenim kot: 1 = brez podrobnosti, 7 = poln podrobnosti). Z določitvijo vrednosti na lestvici so udeleženci opisali tudi občutke pri gledanju in po gledanju posnetka, pogostost razmišljanja in pogovarjanja o dogodku, ocenili so tudi gotovost v lastni spomin.

Postopek

Sedem dni po izpolnjevanju osebnostnih vprašalnikov so si udeleženci ogledali posnetek roparskega napada. Potekal je v manjši amfiteaterski predavalnici, udeleženci so bili razdeljeni na več skupin po 30 oseb. Pred ogledom posnetka so bili udeleženci zaproseni za sodelovanje v raziskavi, ki jo izvaja avtor s skupino kriminalistov. Omenjeno je bilo, da želimo preveriti veljavnost hipotez, ki smo jih postavili pri preiskovanju kaznivega dejanja. Stik med zaporednimi skupinami udeležencev, ki so gledale posnetek, je bil preprečen. Udeleženci so bili naprošeni, da se o videnem ne pogovarjajo med seboj, pojasnjeno jim je bilo, da lahko s pogovori popačijo izvirne spomine. Teden dni po ogledu posnetka so bili udeleženci znova pozvani v predavalnico, v kateri so bili podobni pogoji kot pri ogledu posnetka.

Rezultati

Analiza neodvisnih spremenljivk je pokazala, da vzorec ni povsem homogen, kar je najbolj vidno pri nevroticizmu, kjer opazimo očitno razliko med moškim in ženskim vzorcem (tabela 1). Do podobne ugotovitve sem prišel tudi v predhodni raziskavi (Areh, 2003). Poleg tega raziskave kažejo, da obstajajo občutne razlike med moškimi in ženskami v značilnostih spominske obnove (Richards in Gross, 2000; Schredl in Piel, 2003; Thorne, 2000), zato sem se odločil, da vse analize opravi ločeno na moškem in ženskem vzorcu. Porazdelitev vrednosti spremenljivk odstopa od normalne distribucije (tabela 2), zato sem pri izračunavanju korelacijskih koeficientov uporabil neparametrični Kendallov koeficient (tau-c).

V tabeli 3 je videti, da so bili udeleženci raziskave bistveno manj ekstravertirani, kot je to predvideno z normami vprašalnika EPQ-R. Kaže, da vzorec ni bil najbolj reprezentativen za slovensko populacijo.

Tabela 1. Analiza homogenosti vzorca udeležencev (ANOVA).

Spremenljivki	Izvor variabilnosti	SS	df	MS	F	p
Psihoticizem, spol	Med skupinama	0,94	1	0,94	0,29	0,59
	Znotraj skupin	915,03	278	3,29		
	Skupaj	915,97	279			
Ekstravertnost, spol	Med skupinama	29,14	1	29,14	3,04	0,08
	Znotraj skupin	2662,85	278	9,58		
	Skupaj	2691,97	279			
Nevroticizem, spol	Med skupinama	181,04	1	181,04	16,15	0,00
	Znotraj skupin	3116,33	278	11,21		
	Skupaj	3297,37	279			
Lestvica lažnivosti, spol	Med skupinama	43,05	1	43,05	7,10	0,01
	Znotraj skupin	1685,35	278	6,06		
	Skupaj	1728,40	279			
Točnost obnove, spol	Med skupinama	0,03	1	0,03	5,76	0,02
	Znotraj skupin	1,57	278	0,01		
	Skupaj	1,60	279			
Količina obnove, spol	Med skupinama	0,00	1	0,00	0,01	0,93
	Znotraj skupin	1,92	278	0,01		
	Skupaj	1,92	279			

Tabela 2. Odstopanje vrednosti neodvisnih spremenljivk od normalne distribucije podatkov.

	Spr	As	SE _{As}	Spl	SE _{Spl}	Lillieforjev test		
						Vrednost	df	p
Moški (N = 119)	P	0,47	0,23	-0,01	0,45	0,14	116	0,00
	E	-0,88	0,23	-0,28	0,45	0,21	116	0,00
	N	0,52	0,23	-0,69	0,45	0,16	116	0,00
	TO	0,06	0,23	0,14	0,45	0,04	116	0,20
	KO	-0,77	0,23	1,24	0,45	0,08	116	0,08
Ženske (N = 161)	P	0,71	0,19	0,76	0,38	0,15	159	0,00
	E	-1,29	0,19	0,74	0,38	0,24	159	0,00
	N	0,30	0,19	-0,99	0,38	0,12	159	0,00
	TO	-0,07	0,19	-0,66	0,38	0,07	159	0,03
	KO	-1,40	0,19	3,18	0,38	0,10	159	0,00

Opombe: Spr = spremenljivka; As = asimetričnost; Spl = sploščenost; P = psihotizem; E = ekstravertnost; N = nevroticizem; TO = točnost obnove; KO = količina obnove.

Tabela 3. Razlike med povprečnimi vrednostmi osebnostnih dimenzij vzorca in normami.

	Os. poteza	Mv	SDv	Mn	t	df	p	PR
Moški (N = 119)	P	3,38	1,85	5,32	-11,45	118	0,00	-1,94
	E	8,83	3,24	18,01	-30,90	118	0,00	-9,18
	N	4,19	3,22	5,05	-2,90	118	0,00	-0,86
	L	3,70	2,32	12,69	-42,28	118	0,00	-8,99
Ženske (N = 161)	P	3,26	1,79	4,71	-10,29	160	0,00	-1,45
	E	9,48	2,98	18,95	-40,26	160	0,00	-9,47
	N	5,82	3,44	7,05	-4,54	160	0,00	-1,23
	L	4,49	2,56	11,22	-33,33	160	0,00	-6,73

Opombe: Os. poteza = osebnostna poteza; P = psihotocizem; E = ekstravertnost; N = nevroticizem; L = L lestvica; Mv = M vzorca; SDv = SD vzorca; Mn = norma; PR = povprečna razlika

Zanesljivost čekliste za spominsko obnovo dogodka sem preveril z izračunom Cronbachovega koeficienta α . Dobljena vrednost znaša 0,63. Ker je število postavk v seznamu veliko ($N = 101$), lahko nizko vrednost α pripišemo nizki variabilnosti odgovorov, saj je pri nekaterih postavkah izrazilo prevladoval odgovor »ne vem«.

Tabela 4. Stopnje povezanosti med spremenljivkami v vzorcu žensk ($N = 161$, uporabljen Kendalllov koeficient korelacije tau-c).

	TO ^a	KO ^b	P	E	N	GOT ^c
Psihotocizem (P)	,02	,00				
Ekstravertnost (E)	,12	,17**				
Nevroticizem (N)	-,13*	-,12*				
Lestvica iskrenosti (L)	-,13	-,06				
Ocena jasnosti spomina	,11	,29**	,05	,15*	-,14	,59**
Ocena barvitosti spomina	,01	,19**	,09	,14*	-,06	,27**
Ocena točnosti spomina	,07	,36**	,05	,17*	-,06	,38**
Ocena prisotnost zvočnih podrobnosti	,07	,23**	,07	,13	-,12	,25**
Ocena jakosti občutkov ob gledanju	-,13*	,03	,11	,16*	,09	,02
Ocena jakosti občutkov pri obnovi	-,12	,03	,02	,13	-,07	,09
O dogodku sem razmišljal/a	-,11	-,03	,00	-,01	,02	,06
O dogodku sem se pogovarjal/a	-,07	,08	-,07	,15	,01	,09
Ocena gotovosti v spomin (GOT)	,09	,20**	,05	,04	-,10	
Točnost obnove (TO)		,16**	,02	,12	-,13*	
Količina obnovljenih podatkov (KO)			,00	,17**	-,12*	

^a točnost obnove (TO);

^b količina obnovljenih podatkov (KO)

^c ocena gotovosti v spomin (GOT).

* $p < ,05$. ** $p < ,01$.

Iz tabele 4 je razvidno, da so pri ženskah ugotovljeni korelacijski koeficienti nizki. Najvišja statistično pomembna korelacija med osebnostno potezo (nevroticizem) in točnostjo spominske obnove znaša le $-0,13$. S količino spominske obnove je najbolj povezana ekstravertnost, vendar je tudi v tem primeru korelacija šibka ($r = 0,17$). Opazimo lahko tudi, da so višje subjektivne ocene o podrobnosti lastnega spomina na dogajanje povezane z bolj obsežno spominsko obnovo.

Tabela 5. Korelacije med spremenljivkami v vzorcu moških udeležencev ($N = 119$, uporabljen Kendallov koeficient korelacije tau-c).

	NO ^a	KO ^b	P	E	N	GOT ^c
Psihoticizem (P)	,04	-,02				
Ekstravertnost (E)	,04	,15**				
Nevroticizem (N)	,14	-,12**				
Lestvica iskrenosti (L)	-,11	,09				
Ocena jasnosti spomina	,10*	,25**	,07	,13**	-,08	,48**
Ocena barvitosti spomina	-,01	,15**	-,05	,15**	-,10*	,26**
Ocena točnosti spomina	,07	,31**	,07	,18**	-,06	,34**
Ocena prisotnost zvočnih podrobnosti	,01	,23**	,03	,11*	-,10*	,19**
Ocena jakosti občutkov ob gledanju	,07	,10	,02	,19**	,03	,12
Ocena jakosti občutkov pri obnovi	-,01	,08	,01	,17**	-,07	,14
O dogodku sem razmišljal/a	,02	,05	-,01	,10	,06	,30**
O dogodku sem se pogovarjal/a	,03	,09	-,01	,06	,07	,04
Ocena gotovosti v spomin (GOT)	,15	,18**	,07	,09	-,10	
Točnost obnove (TO)		,12**	,04	,04	,14	
Količina obnovljenih podatkov (KO)			-,02	,15**	-,12**	

^a točnost obnove (TO);

^b količina obnovljenih podatkov (KO)

^c ocena gotovosti v spomin (GOT).

* $p < ,05$. ** $p < ,01$.

Korelacije med spremenljivkami pri moških (v tabeli 5) so še nižje kot pri ženskah. Točnost obnove ni povezana z nobeno od osebnostnih dimenzij. Količina obnove je šibko, a pozitivno povezana z ekstravertnostjo ($r = 0,15$) in negativno z nevroticizmom ($r = -0,12$). Tudi pri moških so subjektivne ocene o podrobnosti spomina na dogodek blago pozitivno povezane s količino obnove in gotovostjo v spomin. Tako kot pri ženskah lahko tudi pri moških opazimo pozitivno šibko povezanost med točnostjo in količino spominske obnove.

Razprava

Z ugotovitvami raziskave ne morem potrditi domneve, da je spominska obnova ekstravertnih oseb bolj točna, kot je obnova introvertnih oseb. Do podobnega zaključka sem prišel tudi v predhodni raziskavi (Areh, 2003). Tako tudi ne morem potrditi prevladujočega mnenja o pozitivni korelaciji med ekstravertnostjo in točnostjo spominske obnove (Aluja idr., 2003; Lieberman, 2000; Ward in Loftus, 1985). Morda lahko vzrok za razliko v ugotovitvah iščemo v uporabi drugačne raziskovalne metodologije, saj so v preteklosti pozitivno povezavo najpogosteje potrjevali z uporabo asociativnih parov (Lieberman, prav tam). Ward in Loftus (prav tam) sta skušala priti do višje ekološke veljavnosti, zato sta uporabila podobno metodo, kot je opisana v tej raziskavi. Pri tem sta ugotovila šibko pozitivno korelacijo, ki statistično ni bila pomembna. Do takšne ugotovitve sem prišel tudi v sam. Avtorji, ki podpirajo domnevo o pozitivni povezanosti ekstravertnosti in točnosti spominskega priklica, se običajno sklicujejo na rezultate raziskav, pri katerih so ugotovili, da visoko osnovno vzburjenje pri introvertih ovira vkodiranje podatkov v dolgoročni spomin (Ward in Loftus, prav tam). Ker je pri ekstravertih nivo osnovnega vzburjenja osrednjega živčevja nižji, te motnje ni, zato je vkodiranje podatkov v dolgoročni spomin bolj učinkovito (Ward in Loftus, prav tam). Vendar če pogledamo nekatere osebnostne lastnosti, ki jih pripisujemo ekstravertnim osebam, npr. impulzivnost, nepremišljenost in ravnodušnost (Eysenck in Eysenck, 1997), potem ni presenetljivo, da korelacija med ekstravertnostjo in točnostjo priklica ni potrjena. Zdi se nelogično, da bi bila spominska obnova impulzivne in nepremišljene osebe točna, prej je pričakovati nasprotno. Kaže, da usmerjenost ekstravertov v njihovo okolje, iskanje novosti in odprtost za nove izkušnje nimajo dovolj velikega vpliva na točnost vkodiranja in kasnejšega priklica podatkov iz epizodičnega spomina.

Drugeče je, če pogledamo odnos med ekstravertnostjo in količino spominske obnove. Pozitivna korelacija med spremenljivkama je potrjena pri moških in pri ženskah ($r_{\text{ženske}} = 0,17, p < 0,01$; $r_{\text{moški}} = 0,15, p < 0,01$). Tudi v predhodni raziskavi sem ugotovil pozitivno, a nekoliko višjo korelacijo med količino priklica in ekstravertnostjo, vendar to velja le za ženske ($r = 0,31; p < 0,05$) (Areh, 2003). V tokratni raziskavi se je domneva delno potrdila tudi pri moških. Menim, da je visoka ekstravertnost povezana z večjo količino spominske obnove zaradi večje zgovornosti ali besedne fluentnosti ter spontanosti, ki sta značilni za ekstravertne osebe.

Kaže, da ekstravertne osebe ocenjujejo lastne spomine kot bolj jasne, barvne in bolj polne podrobnosti. Občutki o dobri kakovosti spomina na dogodek so deloma neupravičeni, saj jih ne spremlja točnejša spominska obnova. Zanimivo je, da pri ekstravertnih osebah subjektivnih ocen o dobri kakovosti spomina ne spremlja večja gotovost v lastni spomin, kar bi bilo pričakovano.

Ocene udeležencev o živahnih, jasnih in podrobnih spominskih predstavah spominjajo na pojav spomina bliskavice, za katerega je značilen tudi intenziven občutek gotovosti v lastni spomin. Iz pregleda korelacijskih koeficientov je razvidno, da obstaja med gotovostjo in spremenljivkami, s katerimi so udeleženci ocenjevali

kakovost spomina, kopica pozitivnih koeficientov korelacije, ki segajo od 0,19 do 0,59 ($p < 0,05$). Do podobnih rezultatov so prišli tudi drugi (Bohanek, Fivush in Walker, 2005). Ti rezultati na prvi pogled podpirajo domnevo o pojavu učinka bliskavice, pri katerem se zaradi intenzivnejšega čustvenega vznemirjenja pojavi precenjevanje točnosti in neupravičena gotovost v lastni spomin (Brown in Kulik, 1977). Vendar pa domneve o pojavu spomina bliskavice ne morem potrditi, ker sem za preverjanje spominske obnove uporabljal strukturirani obrazec (»čekliksto«). Z njim sem ugotavljal pojavljanje popačenih vsebin. Prisotnosti dodanih ali izmišljenih vsebin, ki so tipične za spomin bliskavice, nisem ugotavljal. Pri prosti, nestrukturirani obnovi dogodka bi bil pojav velike količine izmišljenih vsebin pomemben argument v korist domneve o pojavu učinka bliskavice.

Iz tabele 4 je razvidno, da je pri ženskah korelacija med nevroticizmom in točnostjo obnove negativna in nizka ($r = -0,13$ $p < 0,05$). Do šibkih negativnih korelacij med točnostjo obnove in nevroticizmom ali čustveno nestabilnostjo so prišli tudi drugi raziskovalci (Hermans idr., 2004). Vzrokov za negativni vpliv nevroticizma na točnost spominskega priklica je več, nobeden pa ni dokončno potrjen. Glede na to, da so udeleženci raziskave verjeli v resničnost posnetega dogodka, lahko slabe rezultate oseb z višjim nevroticizmom pripišemo anksioznosti. Udeleženci so bili namreč opozorjeni, da sodelujejo pri preverjanju hipotez, povezanih s preiskovanjem resničnega kaznivega dejanja. Zato so se znašli v neke vrste preizkušnji in morda domnevali, da bo njihovo sodelovanje vplivalo na iskanje ter obsodbo storilca. Udeleženci z višjim nevroticizmom so bili zato bolj anksiozni, zaradi višje anksioznosti pa so bili tudi bolj previdni ali zadržani pri odgovarjanju (Sarason, 1972, po Dobson in Markham, 1993). Slednje je lahko vplivalo na znižanje količine navedenih podatkov in posredno na točnost obnove.

Pri moških hipoteze o vplivu nevroticizma na točnost obnove ne morem potrditi (tabela 5). Kaže pa, da nevroticizem pri moških, tako kot pri ženskah, omejuje količino spominskega priklica ($r = -0,12$ $p < 0,01$). Kot razlog za negativno povezavo lahko znova navedem anksioznost, mogoče pa je, da so tudi pri moških na rezultate vplivali občutki ogroženosti zaradi testnih okoliščin, ki jih lahko pričakujemo pri osebah z višjim nevroticizmom (Sarason, 1972, po Dobson in Markham, 1993). Občutki ogroženosti so se lahko pojavili kljub temu, da sem udeležencem poudaril, da ne gre za tekmovalne okoliščine. Negativni vpliv nevroticizma na količino priklicanih informacij lahko pojasnimo tudi z drugimi osebnostnimi lastnostmi, ki so značilne za nevroticizem. Zaskrbljenost, bojzani, pretirana previdnost, neodločenost, občutki krivde, nesproščenost in napetost, torej lastnosti, ki so značilne za nevroticizem (Eysenck in Eysenck, 1997), bi lahko zniževale učinkovitost spominske obnove.

Če primerjamo normativne vrednosti dimenzij vprašalnika EPQ-R z vrednostmi, ki sem jih dobil v raziskavi, lahko opazimo nekaj očitnih razlik (tabela 3). Povprečne vrednosti osebnostnih dimenzij pri vzorcu bistveno odstopajo od norm pri dimenziji ekstravertnost. Pri ženskah je povprečna vrednost na dimenziji ekstravertnost za 9,5 točke nižja od normativne vrednosti, pri moških pa za 9,2 točke. Zaključim lahko, da raziskovalni vzorec odstopa od vrednosti, ki so značilne za normativno

populacijo (študenti, stari od 18 do 20 let). V primerjavi z normativnim vzorcem je v raziskovalnem vzorcu več oseb z višjo introvertnostjo. Če drži, da imajo introvertne osebe manj točen oz. slabši spominski priklic (Aluja idr., 2003; Ward in Loftus, 1985), potem so lahko neizraziti rezultati te raziskave posledica uporabe selekcioniranega vzorca in s tem majhne variabilnosti v osebnostnih lastnostih udeležencev. Kot sem že omenil, korelacija med introvertnostjo in učinkovitostjo spominske obnove še ni potrjena, zato o vplivu selekcioniranosti vzorca ne moremo biti prepričani.

Sklenem lahko, da osebnostne poteze, ki jih merimo z vprašalnikom EPQ-R, nimajo velikega vpliva na spominsko obnovo. Kaže, da je vpliv osebnostnih potez na vkodiranje podatkov v dolgoročni spomin tako majhen, da nad njimi prevlada vpliv situacijskih dejavnikov. Razmišljanje o tem, da zaradi prevlade situacijskih dejavnikov ne moremo zanesljivo napovedovati vedenja in spoznavnih procesov posameznika, se je pojavilo konec 60-ih let prejšnjega stoletja in od takrat se vse bolj potrjuje (Mischel, 2004). Zaključim lahko, da modela, s pomočjo katerega bi zanesljivo napovedovali verodostojnost pričanja očividcev kaznivih dejanj ali nesreč, ni mogoče izdelati. Poleg predhodno omenjenega razloga za to (nezmožnost napovedovanja vedenja in spoznavnih procesov zaradi vpliva nenadzorovanih situacijskih dejavnikov) nas omejuje tudi raziskovalna etika, zaradi katere ne moremo doseči višje ekološke veljavnosti raziskovalnih ugotovitev. Etika namreč upravičeno preprečuje izvedbo raziskav v okoliščinah, v katerih bi v udeležencih inducirali tako intenzivne in kompleksne spoznavne procese, kot se pojavljajo v resničnih okoliščinah.

Literatura

- Aluja, A., Garcia, O. in Garcia, L. F. (2003). Relationships among extraversion, openness to experience and sensation seeking. *Personality and Individual Differences*, 35(3), 671–680.
- Areh, I. (2003). *Veljavnost pričevanja [Validity of eyewitness testimony]*. Neobjavljeno magistrsko delo, Univerza v Ljubljani, Ljubljana, Slovenija [Unpublished masters thesis, University of Ljubljana, Ljubljana, Slovenia].
- Areh, I. (2007). *Vpliv osebnostnih značilnosti na verodostojnost pričanja [An influence of personality characteristics on the reliability of eyewitness testimony]*. Doktorska disertacija, Univerza v Ljubljani, Ljubljana, Slovenija [Unpublished doctoral dissertation, University of Ljubljana, Ljubljana, Slovenia].
- Bartol, C. R. in Bartol, A. M. (2006). *Current perspectives in forensic psychology and criminal justice*. London: Sage.
- Bohanek, J. G., Fivush, R. in Walker, E. (2005). Memories of positive and negative emotional events. *Applied Cognitive Psychology*, 19(1), 51–66.
- Brendle, J. R. in Wenzel, A. (2004). Differentiating between memory and interpretation biases in socially anxious and nonanxious individuals. *Behaviour Research and Therapy*, 42(2), 155–171.
- Brown, R. in Kulik, J. (1977). Flashbulb memories. *Cognition*, 5(1), 73–99.
- Dobson, M. in Markham, R. (1993). Individual differences in anxiety level and eyewitness memory. *Journal of General Psychology*, 119(4), 343–350.

- Eysenck, H. J. (1970). *The structure of human personality* (3. izd.). London: Methuen.
- Eysenck, H. J. (1981). *A model for personality*. New York: Springer.
- Eysenck, H. J. in Eysenck, S. B. G. (1997). *Eysenckove osebnostne lestvice EPQ-R, IVE. Priročnik. [Manual of the Eysenck Personality Scales (EPQ-Revised, IVE)]*. Ljubljana: Center za psihodiagnostična sredstva.
- Freud, S. (1984). Bemerkungen über einen Fall von Zwangsneurose (Rattenmann). V Žižek, S. (ur. in prev.) (1984), *Dve analizi* (str. 111–179). Ljubljana: DDU Univerzum. (Originalno delo objavljeno leta 1909)
- Gilboa-Schechtman, E., Erhard-Weiss, D. in Jeczemien, P. (2002). Interpersonal deficits meet cognitive biases: Memory for facial expressions in depressed and anxious men and women. *Psychiatry Research*, 113(3), 279–293.
- Gomez, R., Gomez, A. in Cooper, A. (2002). Neuroticism and extraversion as predictors of negative and positive emotional information processing: Comparing Eysenck's, Gray's, and Newman's Theories. *European Journal of Personality*, 16(5), 333–350.
- Hergenhahn, B. R. (2001). *An introduction to the history of psychology* (4. izd.). Belmont: Wadsworth Publishing.
- Hermans, D., Broeck, K., Belis, G., Raes, F., Pieters, G. in Eelen, P. (2004). Trauma and autobiographical memory specificity in depressed inpatients. *Behaviour Research and Therapy*, 42(7), 775–789.
- Kverno, K. S. (2000). Trait anxiety influences on judgements of frequency and recall. *Personality and Individual Differences*, 29(3), 395–404.
- Leigland, L. A., Schulz, L. E. in Janowsky, J. S. (2004). Age related changes in emotional memory. *Neurobiology of Aging*, 25(8), 1117–1124.
- Levine, L. J. in Safer, M. A. (2002). Sources of bias in memory for emotions. *Current Directions in Psychological Science*, 11(5), 169–173.
- Libkuman, T. M., Griffith, J., Wines, W. M., Dickel, M. J. in Doty, K. G. (1998). Individual differences in arousal and accessibility to information in memory. *Journal of General Psychology*, 125(4), 367–396.
- Lieberman, M. D. (2000). Introversion and working memory: Central executive differences. *Personality and Individual Differences*, 28(3), 479–486.
- Loftus, E. F. (1996). *Eyewitness testimony*. London: Harvard University Press.
- Mansell, W. (2004). Cognitive psychology and anxiety. *Psychiatry*, 3(4), 6–11.
- Martin, M., Ward, J. C. in Clark, D. M. (1983). Neuroticism and the recall of positive and negative personality information. *Behaviour Research and Therapy*, 21(5), 495–503.
- Mischel, W. (2004). Toward an integrative science of the person. *Annual Review of Psychology*, 55, 1–22.
- Münsterberg, H. (1908). *On the witness stand*. Pridobljeno 10.2. 2008 s spletne strani York University, Toronto, <http://psychclassics.yorku.ca/Munster/Witness/>
- Mogg, K. in Bradley, B. P. (1998). A cognitive-motivational analysis of anxiety. *Behaviour, Research and Therapy*, 36(9), 809–848.
- Myers-Briggs, I. in Myers, P. B. (1980). *Gifts differing: Understanding personality type*. Palo Alto, CA: Davies-Black.
- Narby, D. J., Cutler, B. L. in Penrod, S. (1996). The effects of witness, target, and situational factors on eyewitness identifications. V Sporer, S. L., Malpass, R. in Koehnken, G. (ur.), *Psychological issues in eyewitness identification* (str. 23–52). Mahwah, NY: Lawrence Erlbaum.
- Niedźwieńska, A. (2003). Distortion of autobiographical memories. *Applied Cognitive Psychology*, 17(1), 81–91.

- Reidy, J. (2004). Trait anxiety, trait depression, worry, and memory. *Behaviour Research and Therapy*, 42(8), 937–948.
- Reidy, J. in Richards, A. (1997). Anxiety and memory: a recall bias for threatening words in high anxiety. *Behaviour Research and Therapy*, 35(6), 531–542.
- Richards, J. M. in Gross, J. J. (2000). Emotion regulation and memory: The cognitive costs of keeping one's cool. *Journal of Personality and Social Psychology*, 79(3), 410–424.
- Rohner, J. C. (2002). The time-course of visual threat processing: High threat anxious individuals eventually avert their gaze from angry faces. *Cognition and Emotion*, 16(6), 837–844.
- Ruiz-Caballero, J. in Bermudez, J. (1995). Neuroticism, mood, and retrieval of negative personal memories. *Journal of General Psychology*, 122(1), 29–35.
- Rustig, C. L. (1999). Interactive effects of personality and mood on emotion-congruent memory and judgement. *Journal of Personality and Social Psychology*, 77(5), 1073–1086.
- Rustig, C. L. in Larsen, R. J. (1997). Extraversion, neuroticism, and susceptibility to positive and negative affect: A test of two theoretical models. *Personality and Individual Differences*, 22(5), 200–213.
- Schacter, D. L. (1996). *Searching for memory: The brain, the mind, and the past*. New York: Basic Books.
- Schredl, M. in Piel, E. (2003). Gender differences in dream recall: Data from four representative german samples. *Personality and Individual Differences*, 35(5), 1185–1189.
- Schwartz, S. (1975). Individual differences in cognition: Some relationships between personality and memory. *Journal of Research in Personality*, 9(3), 217–225.
- Shapiro, P. N. in Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin*, 100, 139–156.
- Stopa, L. in Clark, D. M. (2000). Social phobia and interpretation of social events. *Behaviour Research and Therapy*, 38(3), 273–283.
- Terry, W. S. in Burns, J. S. (2001). Anxiety and repression in attention and retention. *Journal of General Psychology*, 128(4), 422–433.
- Thorne, A. (2000). Personal memory telling and personality development. *Personality and Social Psychology Review*, 4(1), 45–56.
- Ward, R. A. in Loftus, E. F. (1985). Eyewitness performance in different psychological types. *The Journal of General Psychology*, 112(2), 191–200.
- Wells, G. L. in Olson, E. A. (2003). Eyewitness testimony. *Annual Review of Psychology*, 54, 277–295.
- Wells, G. L., Memon, A. in Penrod, S. D. (2006). Eyewitness evidence: Improving its probative value. *Psychological Science in the Public Interest*, 7(2), 45–75.
- Wenzel, A., Pinna, K. in Rubin, D. C. (2004). Autobiographical memories of anxiety-related experiences. *Behaviour Research and Therapy*, 42(3), 329–341.
- Yerkes, R. M. in Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, 18(5), 459–482.

Ugotavljanje vzorca stresorjev pri delu učiteljev v povezavi z zadovoljstvom pri delu

*Branko Slivar**
Zavod RS za šolstvo, Ljubljana

Povzetek: Pri delu učiteljev delujejo različni stresorji in povzročajo pri njih delovni stres. Zaradi stresa začne upadati delovno zadovoljstvo, kar vpliva na slabše delovne rezultate, kot so npr. delovna učinkovitost, slabša motivacija in šibkejša pripadnost organizaciji. V naši študiji nas je poleg odnosov in povezav med stresom in zadovoljstvom zanimala predvsem struktura odnosov med stresorji (kako in na kakšen način se različni delovni stresorji kombinirajo in oblikujejo vzorec) in posledice za zadovoljstvo z delom. Pri tem smo uporabili sekvenčno vejno analizo odgovorov (sequential tree analysis – SPSS Answer Tree), ki nam omogoča opis vzorca stresorjev in povezav med njimi v odnosu do različnih posledic njihovega delovanja in vizualni prikaz v drevesni obliki, kar olajša naše razumevanje in vpogled v odnos stresor – posledica. V raziskavo smo vključili 442 osnovnošolskih in 191 gimnazijskih učiteljev. Rezultati so pokazali, da so različni vzorci stresorjev povezani z različnimi stopnjami zadovoljstva z delom in da se vzorec stresorjev, ki delujejo na učitelje in vplivajo na njihovo zadovoljstvo z delom, razlikuje glede na podsistem, v katerem delajo.

Ključne besede: stres, učitelji, vzorci stresorjev, sekvenčna vejna analiza

Examination of the pattern of teachers' work stressors in relation to job satisfaction

Branko Slivar
National Education Institute, Ljubljana, Slovenia

Abstract: Stress is affecting teachers in their daily work and is related to low job satisfaction, low work motivation, low affiliation to organization etc. The study explored not only the relationship between teacher stress and job satisfaction but also the structure of patterns between various teacher's work stressors and particular elements of job satisfaction. In order to develop better understanding of the nature of the stressor experience, a study was undertaken to explore the stressor-job satisfaction relationship using sequential tree analysis. The study included 442 primary school teachers and 191 gymnasium school teachers. Results showed that different stressor patterns were associated with different levels of satisfaction and that there are differences in structure and in patterns of stressors among teachers from different types of schools.

Keywords: stress, teachers, stressor patterns, sequential tree analysis

CC = 3510

* Naslov/Address: Dr. Branko Slivar, Zavod RS za šolstvo, Poljanska 28, Ljubljana, Slovenija, e-mail: branko.slivar@zrss.si

Danes je poklicni stres prepoznan kot eden izmed pomembnih neposrednih in posrednih vzrokov za nezadovoljstvo z delom, anksioznost, depresivno razpoloženje, koronarne srčne bolezni, absentizem ipd. (Jex, 1998). Raziskovalci so ob samem proučevanju stresa seveda identificirali in klasificirali tudi širok razpon stresorjev. Največkrat jih razvrščajo v šest glavnih dejavnikov (Cartwright in Cooper, 1997; Cooper, Dewe in O'Driscoll, 2001): (i) intrinzični dejavniki poklica (delo samo), (ii) vloge posameznika v organizaciji, (iii) medosebni odnosi pri delu, (iv) razvoj kariere, (v) organizacijski dejavniki ter (vi) povezava med delom in zasebnim življenjem. Prvih pet dejavnikov je povezanih s stresorji znotraj delovnega okolja, šesti dejavnik pa predstavlja vire, ki so zunaj delovnega mesta, vendar pa so povezani s poklicnim stresom (npr. družina).

Stres pri učiteljih je oblika poklicnega stresa in številne raziskave s tega področja dokazujejo, da je opravljanje tega poklica stresna dejavnost (Borg in Riding, 1991; Dunham, 1992; Galloway, Panckhurst, Boswell, Boswell in Green, 1987; Jacobsson, Pousette, Thylefors, 2001; Kyriacou in Sutcliffe, 1978; Travers in Cooper, 1996 in 1998). Tudi številni raziskovalci stresa pri učiteljih so proučevali stresorje in jih razvrščali v posamezne skupine. Tako je Kyriacou (1990) razvrstil vire stresa, ki delujejo na učitelje, v šest glavnih skupin: (i) šibka motivacija učencev, (ii) nedisciplina učencev, (iii) slabi delovni pogoji, (iv) časovni pritiski, (v) nizek status učiteljevega poklica, in (vi) konflikti s sodelavci.

Šibka motivacija oziroma pomanjkanje motivacije učencev za učenje in njihov negativen odnos do šole predstavlja že vrsto let v mnogih študijah enega glavnih virov stresa pri učiteljih. Temu najverjetneje botruje na eni stran trud, ki ga učitelji vlagajo v poučevanje takšnih učencev, in na drugi strani majhne možnosti za izboljšanje tega stanja. S šibko motivacijo se povezuje nedisciplina učencev. Slabi delovni pogoji vključujejo neustrezno opremo za izvajanje pouka, neustrezne prostore za osebje, poučevanje na dveh lokacijah ipd. Časovni pritiski na splošni ravni pomenijo, da se od učitelja zahteva, da v kratkem času opravi veliko število nalog. Dejansko je učiteljev tipičen vsakdan sestavljen iz množice nalog, ki morajo biti opravljene do določenega roka (Kyriacou, 1990). Učitelji zaznavajo, da širša družba ne ceni oziroma slabo ceni njihov poklic. To se odraža v nizkih plačah in razpravah o vzgoji in izobraževanju, posebej v medijih. Kritike učiteljevega dela, izpostavljanje napak učiteljev, obtoževanje za povečanje nasilja v družbi ipd. v časopisih in na televiziji nižajo poklicno samospoštovanje in povečujejo stres. Spori med sodelavci prav tako povzročajo stres. Taki spori lahko segajo od nestrinjanja o posameznih strokovnih vprašanjih do sporov, povezanih z vodenjem šole.

J. C. Travers in Cooper (1996) sta v raziskavi stresa na vzorcu 1790 osnovnošolskih in srednješolskih učiteljev s faktorsko analizo identificirala deset faktorjev učiteljevega stresa:

1. interakcija učitelj – učenec (faktor vključuje aktualno vedenje učencev do učiteljev, od besedne do fizične agresivnosti, spoprijemanje s takšnim

- vedenjem, vedenje med učenci in pomanjkanje komunikacije s starši nediscipliniranih otrok);
2. vodenje in organizacija šole (faktor vključuje organizacijsko strukturo in vodenje šole);
 3. preštevilčni razredi (faktor vključuje preveliko število učencev v razredu in preveliko število razredov za posameznega učitelja);
 4. spremembe v vzgoji in izobraževanju (faktor vključuje kurikularne spremembe – novi učni načrti, nove metode in tehnike poučevanja, novi učbeniki itd.);
 5. ocenjevanje učiteljev (faktor vključuje ocenjevanje učiteljevega dela s strani staršev, inšpekcije in ravnateljev);
 6. vodenje (faktor vključuje probleme, povezane z vodenjem predmetnih aktivov in manjših delovnih skupin na šoli);
 7. nizek status (faktor vključuje nizke plače, negativno mnenje družbe, pomanjkanje priložnosti za napredovanje);
 8. nadomeščanje (faktor vključuje pogosta nadomeščanja sodelavcev zaradi njihove odsotnosti);
 9. možnost izgube službe (faktor vključuje možnost prerazporejanja in odpuščenja zaradi zmanjševanja števila vpisanih učencev);
 10. nejasnost učiteljeve vloge (faktor vključuje različne vidike učiteljevih vlog).

Zanimivi so rezultati manjše študije, izvedene na eni od slovenskih osnovnih šol, kjer sta C. Skok in J. Mihalič (1996) na osnovi spremljanja problematike, ki se je obravnavala na konferencah učiteljskega zbora, razdelili stresogene dejavnike v tri večje skupine: področje dela z učenci, področje obremenitev, povezanih s strokovnim delom, in področje odnosov s sodelavci. Področje dela z učenci vključuje agresivno obnašanje učencev, odnos učencev do šolskega dela, nespoštljiv odnos do učitelja ipd. Področje obremenitev, povezanih s strokovnim delom učitelja, vključuje status, ki ga ima ta poklic v družbi, finančno vrednotenje dela, občutek nedokončanosti dela, vsiljen ritem dela, preveč administrativnih opravil, različne poklicne vloge učitelja, občutek nemoči pri reševanju problemov v razredu, občutek preobremenjenosti z delom, občutek nezadostne strokovne usposobljenosti ipd. Področje odnosov s sodelavci vključuje nerazumevanje in nevoščljivost med kolegi, nov položaj učitelja v odnosu do nadrejenih in šolskih oblasti, pomanjkanje povratnih informacij, kritike staršev ipd.

Podobne dejavnike stresa opazimo tudi pri slovenskih osnovnošolskih učiteljih. K. Depolli (1999) je na vzorcu 118 učiteljev s pomočjo vprašalnika odkrila sedem faktorjev: delovna preobremenjenost (časovna stiska, veliko število časovnih rokov v kombinaciji s preveliko količino delovnih nalog), nediscipliniranost učencev, zahtevana samostojnost pri izvedbi učnega procesa, zahtevana kompleksnost dela, pomanjkljiva socialna podpora na delovnem mestu, nezadovoljstvo s poklicnim statusom in kognitivni, emocionalni in vedenjski primanjkljaji učencev. Najbolj obremenjujočo kombinacijo stresogenih faktorjev na delovnem mestu učitelja tvorijo: zahtevana

samostojnost pri izvedbi učnega procesa, nezadovoljstvo s poklicnim statusom in nediscipliniranost učencev. Izmed posameznih stresorjev so učitelji razvrstili na prvih deset mest, kot zelo stresne, naslednje situacije: agresivnost učencev, nevljudnost ali predrznost učencev, disciplinsko težavni razredi, večji vedenjski problemi, odgovornost za varnost učencev (na izletu), visok nivo hrupa v razredu, nizek status učiteljskega poklica, nediscipliniranost učencev, neustrezna politika discipliniranja v šoli in slab odnos do dela pri učencih. Najmanj stresni so: izbira učnih pripomočkov, razlikovanje med manj in bolj pomembno snovjo in priprava zadolžitev za učence. Ugotovljeno je bilo tudi naslednje:

1. za učiteljice je nediscipliniranost učencev pomembnejši izvor stresa kot za učitelje;
2. učitelji slovenščine in matematike (na koncu osmega razreda njihovi učenci opravljajo zunanje preverjanje znanja) občutijo večjo delovno preobremenjenost kot ostali učitelji,
3. učiteljem, ki so mlajši kot 30 oziroma starejši kot 45 let, zahtevana samostojnost pri izvedbi učnega procesa predstavlja manjši izvor stresa kot učiteljem, ki so stari od 30 do 45 let,
4. največ nezadovoljstva s svojim poklicnim statusom občutijo učitelji, ki so na šoli že 16 let ali več.

Slivar (2003) je na vzorcu 191 slovenskih učiteljih na gimnazijah identificiral 10 faktorjev učiteljevega stresa:

1. Trpinčenje (vključuje tiste vzroke stresa, ki jih učitelji doživljajo kot šikaniranja s strani vodstva).
2. Pogoji dela (vključuje vzroke stresa, ki so povezani z neustreznimi pogoji dela na šoli, v prvi vrsti pomanjkanje administrativno-tehnične podpore, učnih pripomočkov in sredstev ter financ).
3. Nedisciplina na šoli (vključuje različne pojavne oblike nediscipline na šoli, kot so poškodovanje in uničevanje šolske opreme, nasilje med učenci in neopravičeno izostajanje).
4. Izvajanje pouka (vključuje predvsem tiste aktivnosti, ki so povezane z uresničevanjem učnega programa pri pouku in so odvisne od časa, ki je opredeljen z omejenim številom ur za posamezni predmet).
5. Odnosi s sodelavci (vključuje neustrezne in negativne medosebne odnose med sodelavci, kot so slaba komunikacija, različni konflikti, pomanjkanje dialoga in slabo delo sodelavcev).
6. Motnje pouka (vključuje tiste motnje pouka, ki so povezane z neprimernim vedenjem učencev v razredu: izzivanje, klepetanje in nesramno vedenje).
7. Vodenje šole (vključuje tiste dejavnike, ki so povezani z načini vodenja: togost, hierarhičnost, možnost participacije zaposlenih pri odločanju, favoriziranje

- nekaterih posameznikov ter slabo pripravljeni in vodeni sestanki).
8. Delo s starši (vključuje dejavnike, ki so povezani s sodelovanjem s starši učencev: pogovori s starši problematičnih učencev, različni konflikti na relaciji starš – učitelj, pomanjkanje podpore staršev, kadar učitelj želi vzpostaviti disciplino v razredu, vodenje roditeljskih sestankov in zahteve staršev po dobrih učnih rezultatih).
 9. Delovna obremenjenost (vključuje nekatere dejavnike, ki učitelja še posebej preobremenjujejo. Taki dogodki so nepredvideno nadomeščanje sodelavcev, prilagajanje pouka učno neuspešnim učencem, iskanje novih didaktičnih prijemov in splošna preobremenjenost z delom).
 10. Nejasnost vloge (vključuje nejasna pričakovanja glede učiteljeve vloge).

Nadaljnje raziskave s področja merjenja stresorjev so pokazale, da ni dovolj samo ugotavljati, kateri stresorji se med seboj povezujejo in kako, ampak je z metodološkega vidika pomembno razmišljati tudi o pomenu različnih stresorjev za posameznika, katere skupine stresorjev so pomembnejše kot druge in katera kombinacija stresorjev vpliva na vedenje posameznika. Ponavadi so raziskovalci z delom povezanih stresorjev uporabljali linearne modele stresorjev in jih med seboj kombinirali (ali računali aritmetično sredino), da bi prišli do skupne vrednosti (skupne aritmetične sredine) stresorjev, ki jih je posameznik doživel (Cooper idr., 2001). Ta skupna aritmetična sredina se nato uporablja kot prediktor za posledice. Tako na primer ni nenavadno za nekatere študije o socialnih vlogah kot virih stresa, da so kombinirali odgovore s področja konfliktnosti vlog, nejasnosti vlog in preobremenjenosti v enotni faktor ali področje, poimenovano stresorji vlog. Tudi znotraj posameznih spremenljivk se kombinirajo postavke, ki predstavljajo posamezne vidike stresorjev. Takšen pristop poenostavlja nadaljnjo statistično analizo z reduciranjem potencialno širokega nabora stresorjev v bolj uporabne in primerne »pakete« ali kategorije (faktorje) ali celo v eno kategorijo ali faktor. Pomembne informacije o posameznih stresorjih se tako skozi proceduro skupne aritmetične sredine skorov izgubijo, saj skupna aritmetična sredina ne daje nobenih indikacij o vzorcu odgovorov, ki sestavljajo skor. Dejansko lahko uporaba skupne aritmetične sredine skorov odvrne pozornost od različnih poti, po katerih je bil ta skupni skor dosežen, in od različnih vzorcev odgovorov, ki jih je posameznik izkazal. Tak pristop lahko pripelje do napačnega zaključka, da so posamezniki, ki so dosegli enak skor, doživeli enako izkušnjo s stresorjem, pri čemer pa so njihove izkušnje lahko dejansko zelo različne.

Dewe in Brook (2000, cit. po Cooper idr., 2001) sta za ilustracijo zgornjih trditev uporabila sekvenčno vejno analizo (angl. *sequential tree analysis*, v nadaljevanju SVA) za raziskovanje različnih vzorcev skorov, doseženih na prediktorski spremenljivki (stresorjih) v primeru, ko so izprašanci dosegli enako skupno aritmetično sredino na kriterijski spremenljivki (napetosti, povezani z delom). Dve skupini izprašancev sta dosegli enak skupni skor na spremenljivki »napetost, povezana z delom« (lestvica napetosti), in sicer 3,23 na petstopenjski lestvici. Skupina stresorjev, ki je bila v tej

študiji raziskana, je vključevala 15 postavk s področja »časovnega pritiska in neneh-
nih občutkov urgentosti«. Pregled vzorca rezultatov, dobljen z uporabo sekvenčne
vejne analize, je pokazal, da je kljub enakemu skupnemu skoru na lestvici napetosti
vzorec stresorjev iz skupine časovnega pritiska, ki so prispevali k napetosti, različen
pri obeh skupinah. Pri skupini 1 sta se v vzorcu pomembnejših stresorjev pojavila
»pritisak ostalih delavcev« in »delovna preobremenjenost«, medtem ko sta bila pri
skupini 2 pomembnejša stresorja »občutek časovne nujnosti« in »pomanjkanje ust-
rezne veljave/uspešnosti/spodbude«. Ta pristop je pokazal na pomembno ugotovitev,
da je kljub enakemu skoru na lestvici napetosti pri obeh skupinah njuna izkušnja
doživljanja tega stresorja različna. To kaže, kako so lahko pri uporabi skupne sredine
pri lestvicah stresorjev kot indeksa doživljanja prezrte razlike v doživljanju stresorjev.
V takem primeru je lahko posledica napačna intervencna pri upravljanju s stresom.
Intervencna, ki bi temeljila na skupnem skoru na lestvici napetosti, povezane z delom,
bo vključevala enake ukrepe za področje časovnega pritiska za obe skupini. To v
osnovi pomeni, da bodo za eno od skupin ti ukrepi neustrezni in posledično manj
učinkoviti ali sploh neučinkoviti. V primeru upoštevanja različnosti vzorcev stre-
sorjev bodo ukrepi pri obeh skupinah, kljub enakemu skupnemu skoru na lestvici
napetosti, različni in prilagojeni odpravljanju specifičnih stresorjev. Prav tako je
zanimivo omeniti, da kljub 15 postavkam, ki tvorijo lestvico napetosti, samo tri
postavke pomembno vplivajo na skor napetosti.

Študije stresa so tudi pokazale, da obstaja tesna povezava med vzroki neza-
dovoljstva z delom in vzroki stresa (Kyriacou in Sutcliffe, 1977). Zadovoljstvo z
delom igra pomembno vlogo pri raziskavah delovnega stresa predvsem kot rezultat
delovanja delovnih stresorjev (Chen, Spector in Jex, 1996). Prav tako je zadovoljstvo
oz. nezadovoljstvo z delom najbolj pogosto raziskana posledica delovanja stresorjev
na delovnem mestu (Cooper idr., 2001). Najbolj konsistenten učinek delovnih stre-
sorjev se kaže v zmanjševanju zadovoljstva z delom. Različne koncepte zadovoljstva
z delom lahko razdelimo v dve večji skupini (Bavendam, 2000). Prvo skupino
sestavljajo koncepti, kjer se pojmuje zadovoljstvo z delom kot večdimenzionalni
pojavi, ki ga sestavlja zadovoljstvo z delom, z delovnimi pogoji, s sodelavci, s plačo,
z napredovanjem ipd. V drugo skupino uvrščamo koncepte, kjer se zadovoljstvo z
delom pojmuje kot enodimenzionalni pojav, ki vključuje splošno zadovoljstvo ali
nezadovoljstvo z delom. V tem primeru zadovoljstvo z delom predstavlja pozitivno
čustveno stanje, ki je posledica delavčeve ocene delovne situacije, in je povezano
z značilnostmi in zahtevami delovnega mesta. Locke (1976, cit. po Konrad, 1994)
definira zadovoljstvo z delom kot posameznikov afektivni odgovor na oceno svojega
dela ali izkušnje, ki so povezane z delom. Podobno definirata Wolverton in Wol-
verton (2002), ki pravita, da se zadovoljstvo z delom pojavi, kadar so posameznikovi
odgovori kos stresorjem, ki zahtevajo rešitev, in kadar doseže optimalno stopnjo
stresa. Na drugi strani pa je nezadovoljstvo z delom manifestacija pritiska, ki je
induciran s pretirano stopnjo stresnosti.

Raziskovalni problem

S študijo smo želeli prispevati k razjasnitvi dveh vprašanj. Prvo je povezano s splošnim vprašanjem o tem, kako različni stresorji oz. kombinacije le-teh učinkujejo na doživljanje posameznikov, v našem primeru na elemente zadovoljstva z delom. Drugo vprašanje pa je povezano z analiziranjem specifičnih stresorjev, ki vplivajo na zadovoljstvo pri delu učiteljev, in ugotavljanjem razlik v vzorcu stresorjev med osnovnošolskimi in gimnazijskimi učitelji. Da bi prišli do odgovorov na obe vprašanji, smo uporabili sekvenčno vejno analizo. Ta analiza sodi k eksploratornim metodam za analizo podatkov in se uporablja za študij odnosov med odvisno spremenljivko (v našem primeru posledice delovanja stresorjev oz. zadovoljstvo z delom) in nizom neodvisnih oz. prediktorskih spremenljivk (stresorji), ki so v interakciji (Trenberth in Dewe, 2006). SVA razširja območje raziskovanja, saj za oblikovanje vzorca stresorjev omogoča vključevanje vseh stresorjev. S tem presežemo težave pri interpretaciji delovanja stresorjev, ki so na osnovi factorske analize združeni v kategorije ali skupine (faktorje) stresorjev in je skupna sredina osrednji predmet analize, kar zamegljuje naravo doživljanja stresorjev (Trenberth in Dewe, 2005). Factorska analiza v takem primeru ustvarja umetne pregrade med različnimi kategorijami stresorjev in tako bolj ali manj vodi k sklepu, da je vpliv posameznih kategorij stresorjev medsebojno neodvisen.

Metoda

Udeleženci

Vprašalnike smo maja 2001 poslali 750 učiteljem razredne in predmetne stopnje na 25 OŠ in 450 učiteljem na 15 gimnazijah. Vrnjenih vprašalnikov iz OŠ je bilo 442 (59 %), iz gimnazij pa 191 vprašalnikov (42 %). V vzorec osnovnošolskih učiteljev je bilo vključenih 89,2 % žensk in 10,8 % moških. Na razredni stopnji je poučevalo 42,5 % in na predmetni 57,5 % učiteljev. Povprečna starost je bila 39,5 let.

Pripomočki

Stresorji na delovnem mestu učiteljev

Za proučevanje mikrostressorjev v učiteljevem delovnem okolju smo uporabili Vprašalnik o šolskih vzrokih stresa pri učiteljih (Slivar, 2003). Vprašalnik vsebuje 67 postavk, ki opisujejo stresorje v šoli. Izprašanec mora vsak stresor oceniti na petstopenjski lestvici, v kolikšni meri mu povzroča stres (situacija se ni pojavila, situacija ne povzroča stresa, situacija povzroča blag stres, situacija povzroča močan stres in situacija povzroča izjemno močan stres). Koeficient notranje konsistentnosti α za celotno lestvico je 0,96.

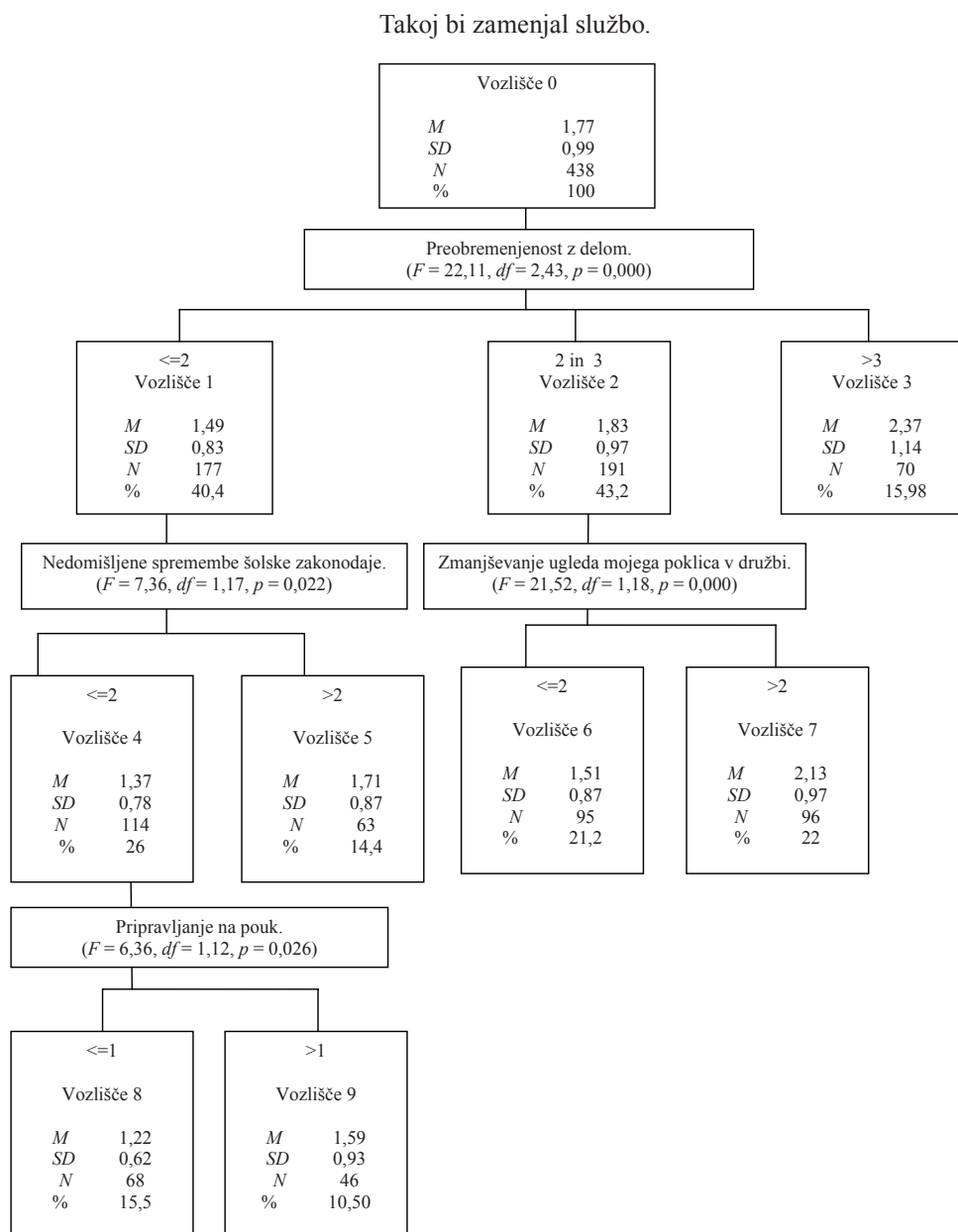
Zadovoljstvo z delom

Za namene naše raziskave smo izhajali iz koncepta enodimenzionalnega pojmovanja zadovoljstva z delom. Za osnovo pri pripravi lestvice smo vzeli postavke z lestvice zadovoljstva, ki jo je uporabil v svoji študiji Wolfgang (1995) in je merila splošno zadovoljstvo. Postavke smo vsebinsko prilagodili našim okoliščinam in izdelali enodimenzionalno lestvico, ki vključuje sedem postavk. Skupni rezultat določimo s seštevanjem ocen na vseh sedmih postavkah. Vprašanec mora na Likartovi štiristopenjski lestvici (1 – se ne strinjam, 2 – se delno ne strinjam, 3 – se delno strinjam, 4 – se strinjam) označiti stopnjo, v kateri se s posamezno trditvijo strinja. Koeficient notranje konsistentnosti α je 0,78.

Za potrebe analize, usmerjene na vzorce stresorjev, ki vplivajo na zadovoljstvo z delom, smo se na osnovi pregleda aritmetičnih sredin odločili uporabiti tri elemente zadovoljstva z delom, ki so se pokazali enaki pri obeh vzorcih učiteljev in so najbolj izstopali po svoji »moči«. To so: (i) »Takoj bi zamenjal službo, če bi imel za to priložnost.« Aritmetična sredina obrnjenih vrednosti pri osnovnošolskih učiteljih je bila 3,22, pri gimnazijskih učiteljih 3,05. (ii) »S svojim poukom sem zadovoljen.« Aritmetična sredina pri osnovnošolskih učiteljih je bila 3,39, pri gimnazijskih učiteljih 3,13. (iii) »Ponovno bi se zaposlil kot učitelj.« Aritmetična sredina pri osnovnošolskih učiteljih je bila 3,37, pri gimnazijskih učiteljih 3,18.

Rezultati

Predpostavili smo, da je vzorec stresorjev, ki vplivajo na izbrane elemente zadovoljstva z delom, drugačen pri osnovnošolskih učiteljih kot pri gimnazijskih učiteljih. Tako smo za proučevanje vzorca stresorjev za posamezne skupine učiteljev in njihov vpliv na tri izbrane elemente zadovoljstva z delom uporabili SVA. Rezultat SVA je drevo, ki prikazuje kako vzorec prediktorskih spremenljivk različno napoveduje odvisne spremenljivke. SVA razdeli vzorec zaporedno v homogene skupine. Najboljša prediktorska spremenljivka odvisne spremenljivke se določi na osnovi F -testa in v drevesu predstavlja prvo vejo. Nato na osnovi druge pomembne prediktorske spremenljivke razdeli vsako homogeno skupino v podskupine, nato po istem principu razdeli te podskupine v še manjše podskupine (veje), do zadnje še statistično pomembne prediktorske spremenljivke. Ker SVA temelji na razdelitvi vrednosti posameznih lestvic, postanejo posamezne postavke stresorjev mera analize. V našem primeru smo za preprečitev predimenzioniranja drevesa (zmanjšanja prevelike razvejanosti) uporabili konzervativni pristop. Opredelili smo dve vrednosti, s katerima smo določili obseg drevesa in velikost vzorca v posameznih drevesnih vozliščih. V primerih, kjer je bila pomembnost F vrednosti manjša od določene (v našem primeru je to 5-odstotna stopnja tveganja) in kjer je bila velikost vzorca v posameznem vozlišču manjša od 30 pri osnovnošolskem vzorcu in manjša od 10 pri gimnazijskem vzorcu, nismo nadaljevali z analizo v tej veji.



Slika 1. Sekvenčna drevesna analiza povezav med "Takoj bi zamenjal službo, če bi imel za to priložnost." in delovnimi stresorji pri osnovnošolskih učiteljih.

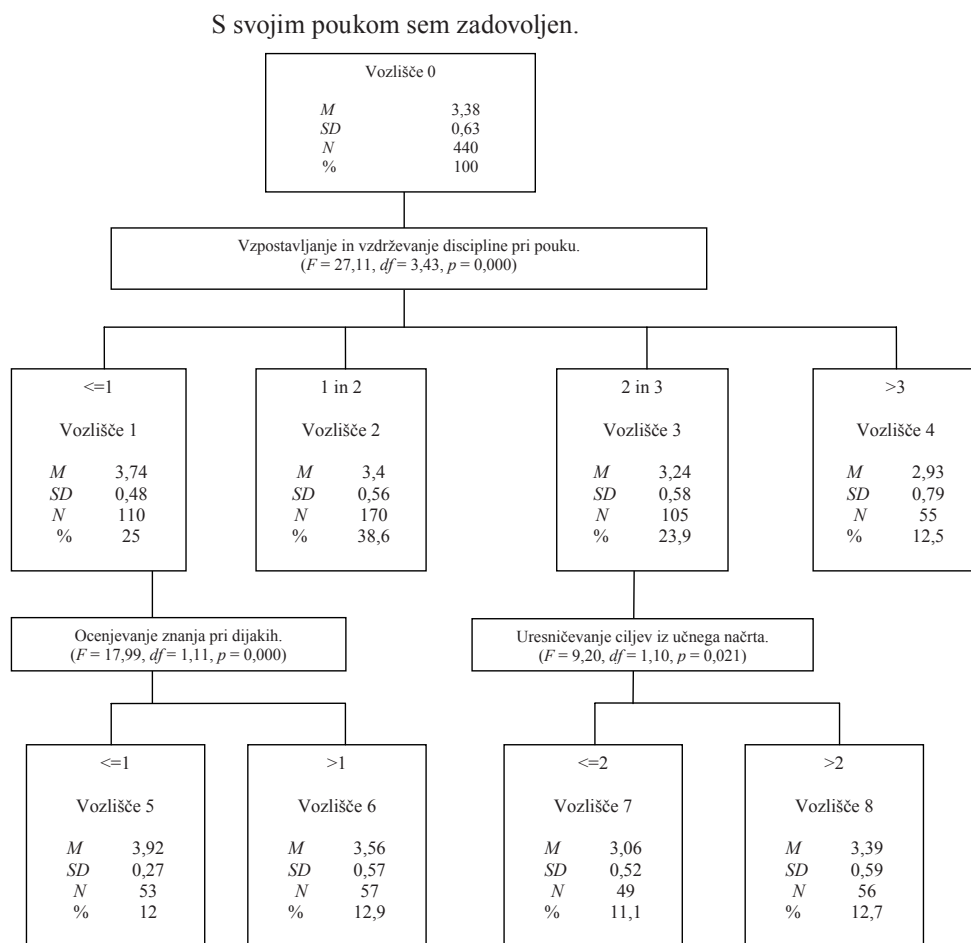
Vzorec stresorjev pri osnovnošolskih učiteljih je prikazan na slikah od 1 do 3 za vsak element zadovoljstva z delom posebej; vzorec stresorjev za gimnazijske učitelje pa je prikazan na slikah od 4 do 6. Na vrhu drevesa (1. raven) vsake slike

je v oknu prikazan povzetek statistike za vsako posledico delovanja stresorjev za celoten vzorec. Nad posameznimi vozlišči, ki predstavljajo homogene skupine, so posamezne stopnje vpliva posameznega stresorja (lestvica od 1 do 4): 1 – situacija ne povzroča stresa, 2 – situacija povzroča blag stres, 3 – situacija povzroča močan stres in 4 – situacija povzroča izjemen stres. Podatki v vsakem vozlišču so: aritmetična sredina posamezne posledice, standardni odklon, število izprašancev v tem vozlišču in odstotek, ki ga skupina predstavlja v celotnem vzorcu izprašancev.

Slika 1 prikazuje povezave med elementom zadovoljstva »*Takoj bi zamenjal službo, če bi imel za to priložnost*« in nekaterimi delovnimi stresorji pri osnovnošolskih učiteljih. Želja po zamenjavi službe narašča v drevesni strukturi od leve proti desni. Preobremenjenost z delom je prvi stresor, ki razlikuje med šibko in močnejšo težnjo po zamenjavi službe. Iz drevesa je jasno razvidno, da so različne stopnje težnje po zamenjavi službe povezane z različnim vzorcem stresorjev. Če pogledamo na desno stran drevesa, ugotovimo, da je močnejša težnja po zamenjavi službe povezana z večjo preobremenjenostjo z delom (2,37). Na levi strani drevesa pa lahko opazujemo vzorec stresorjev, ki ga sestavljajo manjša preobremenjenost z delom, nedomišljene spremembe šolske zakonodaje in pripravljanje na pouk. Aritmetična sredina težnje po zamenjavi službe za to vozlišče narašča, vendar je naraščanje majhno od 1,37 do 1,71 na drugi ravni. Tretji vzorec stresorjev je prikazan v srednjem delu drevesa, kjer je preobremenjenost z delom v kombinaciji z zmanjševanjem ugleda poklica v družbi.

Če rezultate, prikazane na sliki 2, kjer je element zadovoljstva »*S svojim poukom sem zadovoljen*«, beremo od leve proti desni, nam kažejo, da je prvi stresor, ki najbolj diskriminira med učitelji, ki so s svojim poukom zadovoljni, vzpostavljanje in vzdrževanje discipline pri pouku. Nadaljnja analiza druge polovice drevesa kaže, da se zadovoljstvo s poukom včasih zmanjšuje, kadar se vzpostavljanje in vzdrževanje discipline pri pouku kombinira z uresničevanjem ciljev iz učnega načrta (3,39). Leva stran drevesa pa sugerira, da kombinacija stresorjev *vzpostavljanje in vzdrževanje discipline* in *ocenjevanje znanja* v manjši meri zmanjšuje zadovoljstvo s poukom. Takšen rezultat odpira vprašanje ali so nekateri stresorji bolj močni kot drugi oz. ali je moč stresorjev odvisna od njihove kombinacije z drugimi stresorji. Seveda se ta dva pogoja med seboj ne izključujeta.

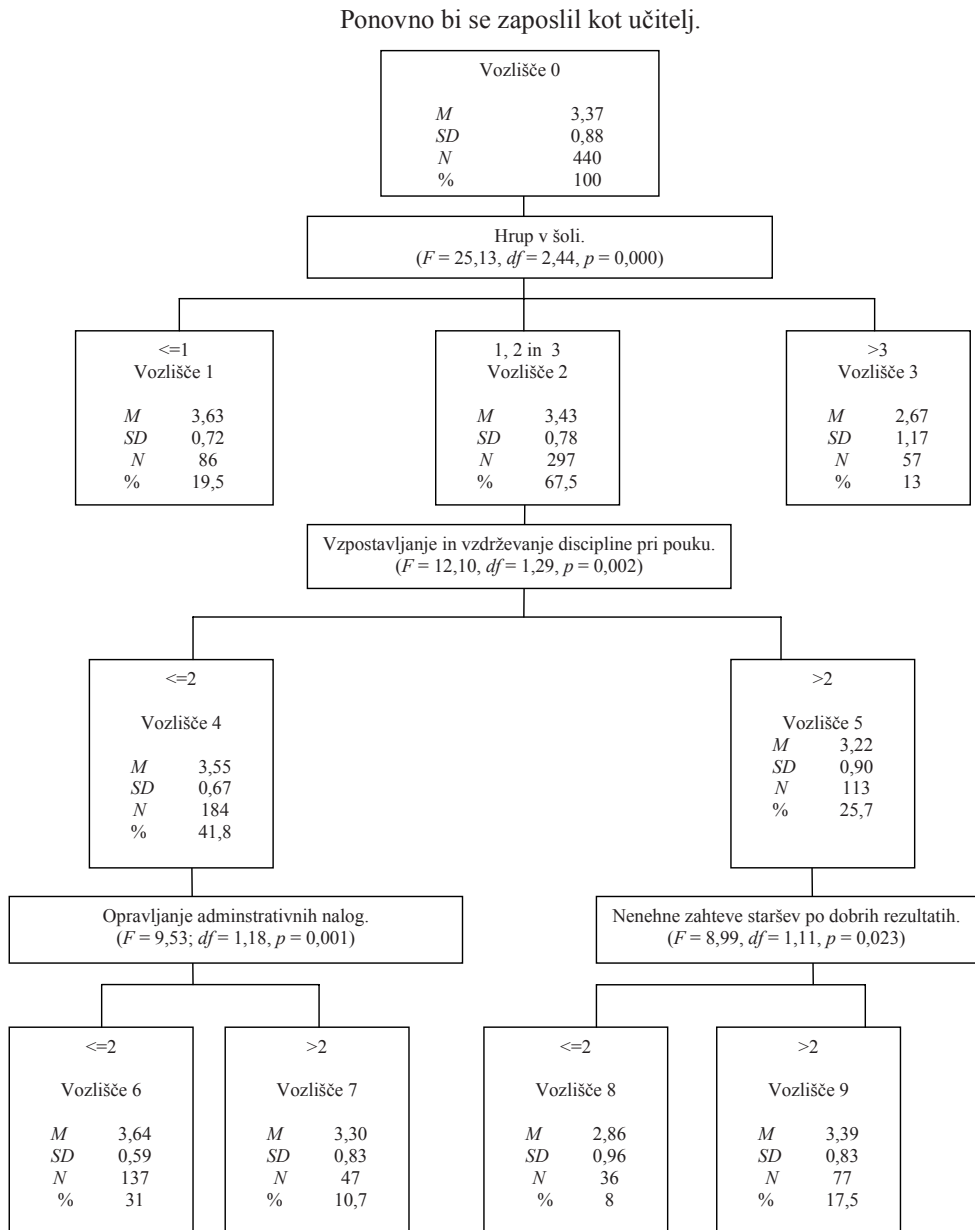
Slika 3 predstavlja rezultate za odvisno spremenljivko »*Ponovno bi se zaposlil kot učitelj*«. Stresor *hrup v šoli* najprej diskriminira med tistimi učitelji, ki bi se ponovno zaposlili kot učitelji, in tistimi, ki so manj prepričani v to. Če pregledamo edino vejo, ki je v drevesu, ugotovimo, da kombinacija stresorjev *hrup v šoli* (3,43), *vzpostavljanje in vzdrževanje discipline* (3,22) in *nenehne zahteve staršev po dobrih rezultatih* (2,86) povečuje verjetnost, da se izprašanci v teh vozliščih ne bi ponovno zaposlili kot učitelji. Tudi to drevo ponovno ilustrira, kako kombinacija različnih vrst stresorjev lahko ojača posamezne elemente zadovoljstva oz. nezadovoljstva z delom. Prav tako drevo kaže, da je stopnja posameznega elementa zadovoljstva z delom lahko odvisna tako od posamičnega stresorja kot od zapletene kombinacije različnih stresorjev.



Slika 2. Sekvenčna drevesna analiza povezav med "S svojim poukom sem zadovoljen." in delovnimi stresorji pri osnovnošolskih učiteljih.

Vejna analiza pri gimnazijskih učiteljih pokaže nekoliko drugačno sliko glede vzorca stresorjev, ki so povezani s posameznimi elementi zadovoljstva z delom.

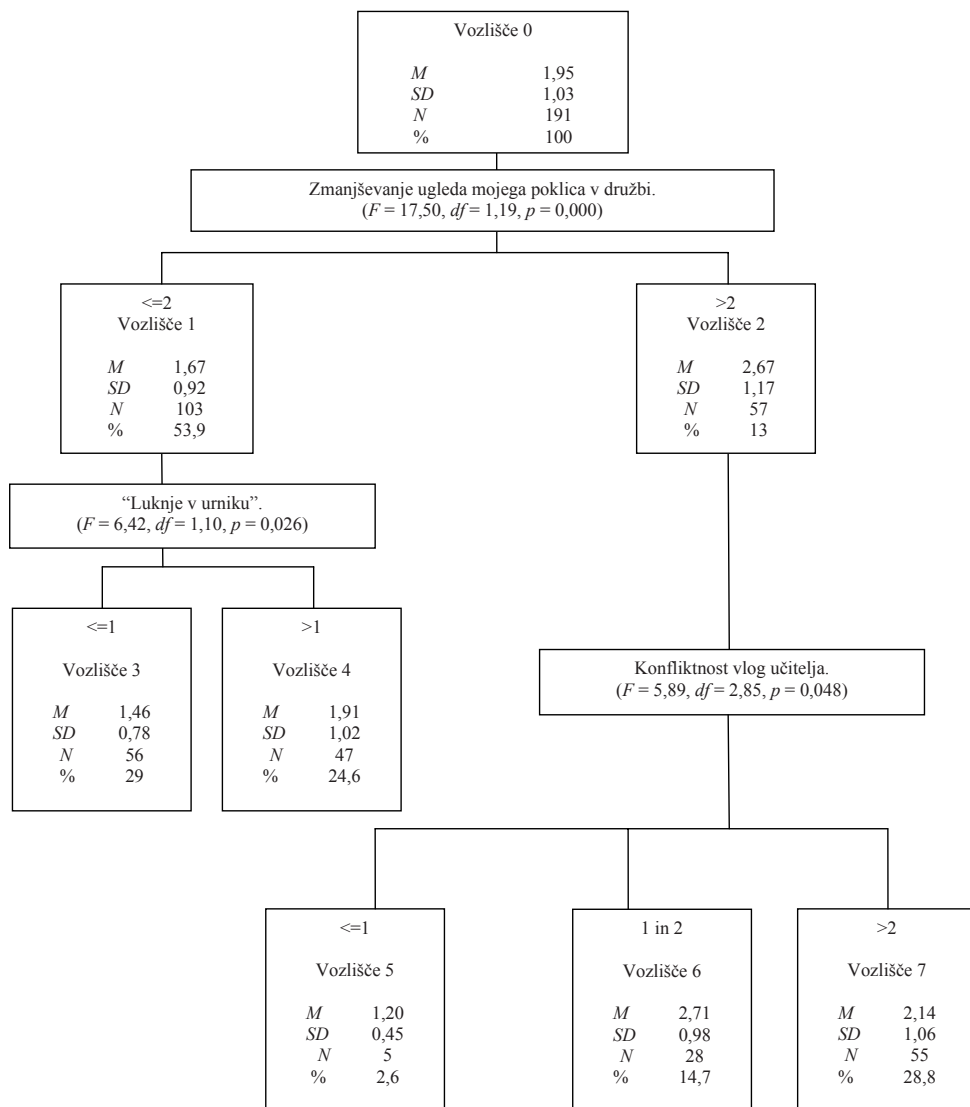
Slika 4 prikazuje povezave med elementom zadovoljstva »*Takoj bi zamenjal službo, če bi imel za to priložnost*« in nekaterimi delovnimi stresorji pri gimnazijskih učiteljih. Želja po zamenjavi službe narašča v drevesni strukturi od leve proti desni. Zmanjševanje ugleda mojega poklica v družbi je prvi stresor, ki razlikuje med šibko in močnejšo težnjo po zamenjavi službe. Iz drevesa je, tako kot pri osnovnošolskih učiteljih, jasno razvidno, da so različne stopnje težnje po zamenjavi službe povezane z različnim vzorcem stresorjev. Če pogledamo na desno stran drevesa, ugotovimo, da je močnejša težnja po zamenjavi službe povezana z večjim zmanjševanjem ugleda poklica (2,27) in konfliktnostjo vlog učitelja (2,14). Aritmetična sredina težnje po zamenjavi službe za to vozlišče narašča, in sicer od 1,20 do 2,14 na drugi ravni. Na



Slika 3. Sekvenčna drevesna analiza povezav med "Ponovno bi se zaposlil kot učitelj." in delovnimi stresorji pri osnovnošolskih učiteljih.

levi strani drevesa pa lahko opazujemo vzorec stresorjev, ki ga sestavljajo manjše zmanjševanje ugleda in luknje v urniku. Aritmetična sredina težnje po zamenjavi službe za to vozlišče narašča, vendar je naraščanje majhno od 1,46 do 1,91 na drugi ravni.

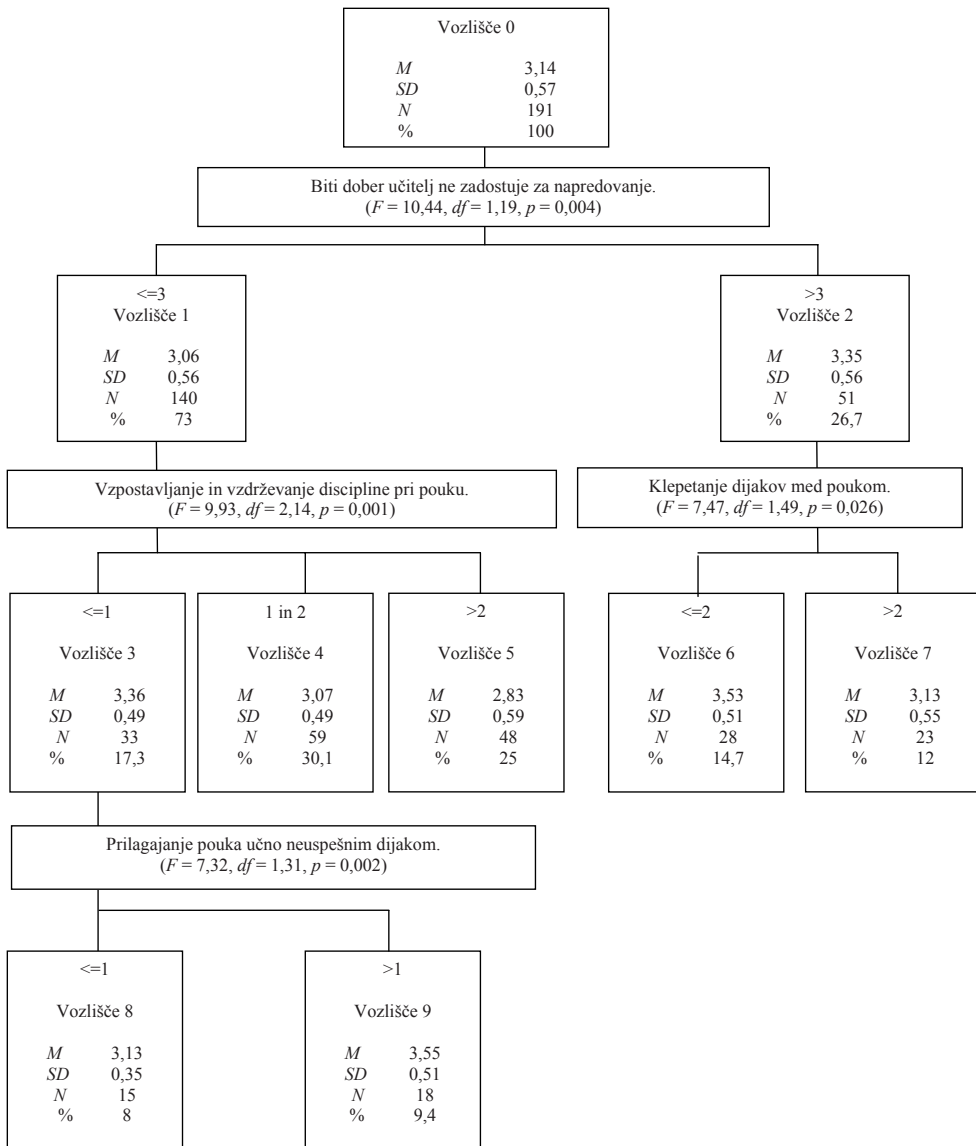
Takoj bi zamenjal službo, če bi imel za to priložnost.



Slika 4. Sekvenčna drevesna analiza povezav med "Takoj bi zamenjal službo, če bi imel za to priložnost." in delovnimi stresorji pri gimnazijskih učiteljih.

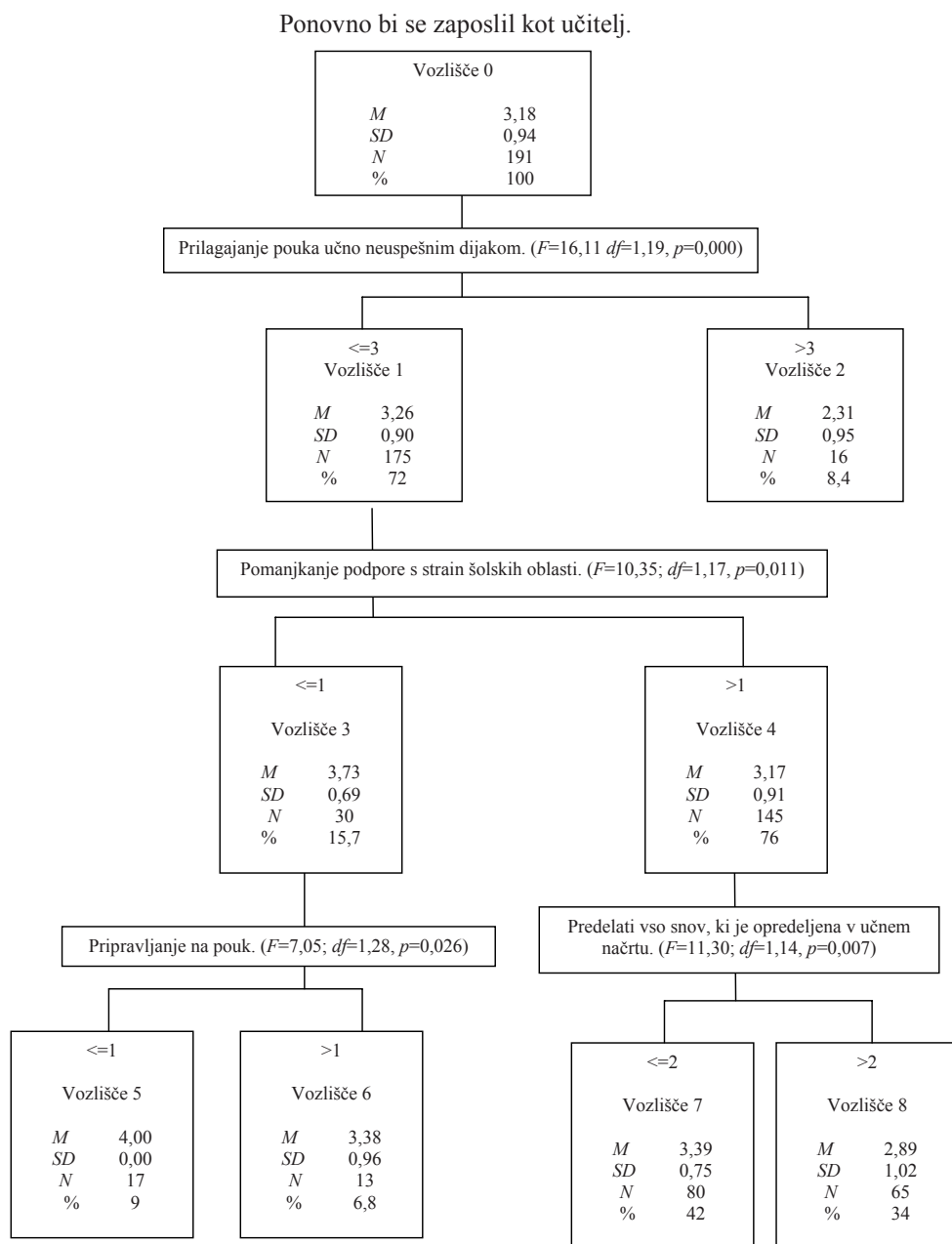
Slika 5 predstavlja rezultate za odvisno spremenljivko »S svojim poukom sem zadovoljen«. Dejstvo, da za napredovanje ne zadostuje biti dober učitelj, je prvi stresor, ki najprej diskriminira med tistimi učitelji, ki so s svojim poukom zadovoljni. Desna stran drevesa kaže, da je ta stresor bolj ogrožujoč za tiste učitelje, ki so svojim poukom zadovoljni (3,35). Ta stresor se pojavlja v kombinaciji s klepetanjem dijakov

S svojim poukom sem zadovoljen.



Slika 5. Sekvenčna drevesna analiza povezav med "S svojim poukom sem zadovoljen." in delovnimi stresorji pri gimnazijskih učiteljih.

med poukom. Na levi strani so učitelji, ki so s svojim poukom manj zadovoljni (3,05), vendar je zanje dejstvo *biti dober učitelj ne zadostuje za napredovanje* manj stresno kot za skupino učiteljev na desni strani drevesa. Pri skupini učiteljev, ki so manj zadovoljni s svojim poukom, se pojavlja na drugem nivoju še stresor *vzpostavljanje*



Slika 6. Sekvenčna drevesna analiza povezav med “Ponovno bi se zaposlil kot učitelj.” in delovnimi stresorji pri gimnazijskih učiteljih.

in vzdrževanje discipline pri pouku in na tretjem nivoju stresor prilagajanje pouka učno neuspešnim dijakom.

In na koncu slika 6. Ta predstavlja rezultate za odvisno spremenljivko »*Ponovno bi se zaposlil kot učitelj*«. Stresor *prilagajanje pouka učno neuspešnim dijakom* najprej diskriminira med tistimi učitelji, ki bi se ponovno zaposlili kot učitelji, in tistimi, ki so manj prepričani v to. Če pregledamo edino vejo, ki je v drevesu, ugotovimo, da kombinacija stresorjev *prilagajanje pouka učno neuspešnim dijakom* (3,26), *pomanjkanje podpore s strani šolskih oblasti* (3,16) in *predelati vsi snovi, ki je v učnem načrtu* (2,89) povečuje verjetnost, da se izprašanci v teh vozliščih ne bi ponovno zaposlili kot učitelji.

Tudi to drevo ponovno ilustrira, kako kombinacija različnih vrst stresorjev lahko ojača posamezne elemente zadovoljstva oz. nezadovoljstva z delom. Prav tako drevo kaže, da je stopnja posameznega elementa zadovoljstva z delom lahko odvisna tako od posamičnega stresorja kot od zapletene kombinacije različnih stresorjev.

Razprava

Naša raziskava je z uporabo SVA pokazala na pomembnost vpliva različnih šolskih stresorjev na zadovoljstvo z delom pri učiteljih osnovnih šol in gimnazij. V raziskavi smo bili usmerjeni predvsem na ugotavljanje, kateri so vzorci stresorjev, ki se pojavljajo pri obeh skupinah učiteljev, in kakšen je njihov vpliv na zadovoljstvo z delom. Rezultati kažejo, da se glede na skupino učiteljev pojavljajo razlike v vzorcih stresorjev, povezanih z zadovoljstvom z delom. To se sklada tudi z ugotovitvami, ki jih je podal Kyriacou (1990), ki je na osnovi primerjav različnih študij o vzrokih stresa pri učiteljih ugotovil, da so lahko za posamezne skupine učiteljev vzroki stresa različni. Rezultati pri osnovnošolskih in gimnazijskih učiteljih kažejo, da so različne vrste stresorjev povezane z različnimi elementi zadovoljstva z delom. Na stopnjo zadovoljstva z delom pri učiteljih iz osnovnih šol vplivajo kot najmočnejši stresorji predvsem preobremenjenost z delom, vzpostavljjanje in vzdrževanje discipline pri pouku in hrup v šoli. Pri gimnazijskih učiteljih pa so to zmanjševanje ugleda učiteljskega poklica v družbi, biti dober učitelj ne zadostuje za napredovanje in prilagajanje pouka učno neuspešnim dijakom. Za osnovnošolske učitelje je težnja po zamenjavi službe najbolj povezana z občutkom preobremenjenosti z delom. Občutek preobremenjenosti z delom nastane kot posledica neustrezne (prevelike) količine dela, ki ga mora nekdo opraviti pod časovnim pritiskom oz. v določenem času in se povezuje z nezadovoljstvom, anksioznostjo, depresijo in delovno neučinkovitostjo (Cooper idr., 2001). Za učiteljevo delo pa je poleg časovnega pritiska značilna tudi fragmentacija nalog (Esteve, 1990), ki še bolj povečuje občutke preobremenjenosti. Učitelji morajo namreč opraviti veliko število posameznih aktivnosti (fragmentov) hkrati ali na različnih ravneh (vzdrževanje discipline v razredu, vendar na dobrohoten ali celo ljubeč način, individualiziranje pouka, vzdrževanje pozitivne razredne klime, načrtovanje dela, preverjanje in ocenjevanje ipd.).

Medtem ko je pri osnovnošolskih učiteljih stresor *preobremenjenost na delovnem mestu* najbolj povezan s težnjo po zamenjavi službe oz. nanjo najbolj vpliva, se pri gimnazijskih učiteljih s težnjo po zamenjavi službe najbolj povezuje stresor

zmanjševanje ugleda poklica v družbi. Ta stresor se pojavlja tudi pri osnovnošolskih učiteljih, vendar v kombinaciji s stresorjem preobremenjenosti, s katerim tvori vzorec stresorjev, ki vplivajo na težnjo po zamenjavi službe. Zmanjševanja ugleda poklica v družbi je stresni proces, ki ga Esteve (1990) uvršča med najpomembnejše sekundarne dejavnike, ki povzročajo nastanek občutkov nemoči. Ti dejavniki imajo izrazito negativen učinek na učiteljevo samopodobo, na učiteljeva stališča do poklica in lahko sprožajo identitetno krizo. Pri gimnazijskih učiteljih se stresor *zmanjševanje ugleda poklica* pri svojem vplivu na težnjo po zamenjavi službe povezuje v vzorec z drugim stresorjem, in sicer s konfliktnostjo vlog učitelja. Konflikt vlog nastane zaradi nasprotujočih oziroma neuskklajenih zahtev do posameznika na delovnem mestu: mora delati stvari, ki jih noče ali pa ne sodijo v njegov delokrog ali pristojnost. Študije so potrdile zvezo med konfliktom vlog in visoko stopnjo stresa. Pri učiteljih se omenjajo naslednji primeri konfliktov vlog (Travers in Cooper, 1996):

1. Učitelj želi opravljati delo drugače, kot je navedeno v opisu del in nalog.
2. Ravnatelj zahteva od učitelja, da poučuje predmet, za katerega ta ni strokovno usposobljen.
3. Učitelj opravlja več vlog, ki si med seboj nasprotujejo, npr. hkrati je ocenjevalec znanja in svetovalec, zaupnik.

(Ne)zadovoljstvo s poukom, kot poseben element zadovoljstva z delom, se pri osnovnošolskih učiteljih povezuje z vzpostavljanjem in vzdrževanjem discipline pri pouku. Bolj se učitelji ukvarjajo z disciplino pri pouku, večja je verjetnost, da bodo nezadovoljni s svojim poukom. Proces vzdrževanja discipline med poukom je za učitelje stresor in je kot dejavnost usmerjena na preprečevanje oz. obvladovanje in nevtraliziranje različnih neprimernih oblik vedenja učencev v razredu med poukom, kot so npr. izzivanje (provociranje) učencev, klepetanje, nesramno (drzno) vedenje. Takšne oblike vedenja učencev pa predstavljajo enega pomembnejših dejavnikov stresa pri učiteljih v Nemčiji, Avstriji, Angliji in Švici (Rudow, 1999). Pri osnovnošolskih učiteljih se vzpostavljanje discipline skupaj z uresničevanjem ciljev iz učnega načrta povezuje v vzorec stresorjev, ki lahko ojačajo občutek nezadovoljstva s poukom. Ukvarjanje z disciplino učencev pri pouku, ki med drugim zahteva dodaten didaktičen napor učitelja in več časa, lahko ogroža doseganje oz. uresničevanje ciljev iz učnega načrta, kar se seveda odraža v nezadovoljstvu učitelja s svojim poukom. Pri gimnazijskih učiteljih je slika nekoliko drugačna. Rezultati kažejo, da dejstvo *biti dober učitelj ne zadostuje za napredovanje*, predstavlja pomemben stresor predvsem za tiste učitelje, ki so sicer zadovoljni s svojim poukom. Na zanimiv pojav pozitivne povezanosti med stopnjo stresa in zadovoljstva z delom pa opozarja Kyriacou (1987). Nekateri učitelji, ki poročajo o visoki stopnji stresa, poročajo tudi o visoki stopnji zadovoljstva z delom, vendar samo takrat, kadar se zadovoljstvo z delom pojavlja v moderirajoči vlogi. Do podobnih ugotovitev sta prišla tudi Cox in Brockley (1984). Menita, da ima zadovoljstvo z delom pri učiteljih, podobno kot socialna podpora, vlogo psihološkega blažilca, ki zmanjšuje škodljive učinke poklicnega stresa. Z napredovanjem na delovnem mestu so povezana dodatna finančna sredstva,

povečan ugled posameznika v njegovem okolju in novi izzivi. Zato pomanjkanje priložnosti za napredovanje ali onemogočanje napredovanja predstavljata pomemben vir nezadovoljstva in stresa pri delavcih (Jewell, 1988, cit. po Cooper idr., 2001). Podobno tudi drugi avtorji, kot npr. Häbler in Kunz (1985), Pieren in Schärer (1992) ter Wulk (1988) (vsi cit. po Rudowu, 1999), govorijo o pomanjkanju priložnosti za napredovanje kot o enem izmed poklicnih stresorjev pri učiteljih. Pri gimnazijskih učiteljih se ta stresor pojavlja v kombinaciji s stresorjema *klepetanje dijakov med poukom* in *vzpostavlanje in vzdrževanje discipline pri pouku*, kar oboje še povečuje občutek nezadovoljstva.

Na zmanjševanje težnje po ponovni zaposlitvi kot učitelj ima pri osnovnošolskih učiteljih najpomembnejši vpliv stresor *hrup v šoli*. Hrup v šoli in razredu povzroča fizično obremenitev, saj v času odmorov dosega vrednost do 93 decibelov (Rudow, 1999), v razredu pa v razponu od 50 do 79 decibelov (Betul, 2000). Za primerjavo: glasen razgovor dosega 40 dB, hrup vrtalca 80 dB, hrup letala z razdalje 4 m pa 120 dB (Pečjak, 1977). Tudi drugi avtorji, kot npr. Scheuch idr. (1978), Leuschner (1979), Wulk (1988), Knothe idr. (1991) (vsi cit. po Rudowu, 1999) prištevajo hrup k pomembnim dejavnikom stresa pri učiteljih.

V kombinaciji s hrupom v šoli se, kot stresor, povezuje tudi vzpostavlanje in vzdrževanje discipline. To kaže na pomembnost tega stresorja, saj se pojavlja v različnih kombinacijah stresorjev, ki vplivajo na posamezne elemente zadovoljstva z delom. Pri gimnazijskih učiteljih je kombinacija stresorjev, ki vplivajo na težnjo po ponovni zaposlitvi kot učitelj, drugačna. Pri njih je v ospredju stresor *prilagajanje pouka učno neuspešnim dijakom*, ki se kombinira z pomanjkanjem podpore s strani šolskih oblasti. Učitelji morajo vsakodnevno prilagajati svoje didaktične pristope različno sposobnim dijakom in med njimi tudi manj uspešnim. Izvajanje takšnega načela pa je težko uresničljivo, posebej še ob upoštevanju normativa, ki določa, naj bo v oddelku 30 dijakov. Tudi sicer se je pri naših učiteljih v gimnazijah kot pomemben stresor pojavilo prav preveliko število dijakov v oddelku (Slivar, 2003). O problematiki prevelikih oddelkov (30 in več dijakov) poročajo tudi drugi raziskovalci iz Nemčije, Avstrije, Anglije in Švice: Schuh (1962), Knight-Wegenstein (1973), Merz (1979), Häbler & Kunz (1985), Urban (1985) (vsi cit. po Rudow, 1999).

Študija je pokazala, kako lahko kombinacija različnih vrst stresorjev ojača posamezne elemente zadovoljstva oz. nezadovoljstva z delom pri učiteljih ter odvisnost stopnje posameznega elementa zadovoljstva z delom tako od posamičnega stresorja kot od zapletene kombinacije različnih stresorjev. Ob teh rezultatih se pojavlja tudi vprašanje, zakaj razlike v vzorcu stresorjev, ki delujejo na enake elemente zadovoljstva, pri osnovnošolskih in gimnazijskih učiteljih. Razlogi so zelo zapleteni in jih s to študijo nismo neposredno ugotavljali. Na osnovi poznavanja dejstev lahko predpostavljamo, da so razlike v vzorcu stresorjev med obema skupinama učiteljev povezane z razliko med učitelji (npr. stopnjo izobrazbe), med učenci (npr. stopnjo intelektualnega in emocionalnega razvoja) ter razliko v sistemu organizacije (npr. različno zakonodajo - pravilniki).

Rezultati študije nakazujejo tudi potrebo po različnih pristopih pri upravljanju stresa za osnovnošolske ali gimnazijske učitelje. Razumevanje vzrokov stresa lahko

predstavlja dobro osnovo za razvijanje strategij, ki jih lahko učitelji uporabljajo pri zmanjševanju individualnega stresa in za spreminjanje prakse v organizaciji in vodenju šole. Predvsem pa je pomembno kontinuirano:

- a) ozaveščanje tako širše javnosti kot učiteljev in vodstev šol o prisotnosti stresa pri učiteljih in o njegovih posledicah,
- b) spodbujanje zavedanja, da delovni stres ni niti problem niti šibkost posameznika, ampak pojav, ki zadeva celotno šolo, in
- c) opozarjanje na možnost pojavljanja stresa oziroma stresnih situacij, tako da to zavedanje postane sestavni del vodenja in upravljanja šole.

Prav tako bi bilo potrebno spodbujati in omogočati še več raziskav s področja stresa pri slovenskih učiteljih, saj je to področje v primerjavi z ostalimi evropskimi državami zanemarjeno.

Raziskava ima tudi svoje omejitve, predvsem majhen vzorec v nekaterih voziščih in ustreznost prilagoditve konzervativnega pristopa pri opredelitvi kriterijev za zmanjšanje prevelike razvejanosti drevesa.

Literatura

- Bavendam, J. (2000). *Managing job satisfaction*. Pridobljeno maja 2007 s spletne strani: <http://www.employeesatisfaction.com>
- Betul, A. (2000). Concentration, noise and anxiety. V *STAR abstracts, 21st International Conference of the Stress and Anxiety Research Society (STAR), July 20–22, 2000, Bratislava, Slovakia*. Bratislava: Stimul-Centrum informatiky a vzdelávania FIF UK.
- Borg, M. G. in Riding, R. J. (1991). Occupational stress and satisfaction in teaching. *British Educational Research Journal*, 17(3), 263–282.
- Cartwright, S. in Cooper, L. C. (1997). *Managing workplace stress*. Thousand Oaks, CA: Sage.
- Chen, P., Spector, P. in Jex, S. (1996). Effects of manipulated job stressors and job attitude on perceived job conditions: A simulation. V S. L. Sauter, in L. R. Murphy (ur.), *Organizational risk factors for job stress* (str. 341–356). Washington: American Psychological Association.
- Cooper, L. C., Dewe, P. J. in O’Driscoll, M. P. (2001). *Organizational stress, a review and critique of theory, research, and applications*. Thousand Oaks, CA: Sage.
- Cox, T. in Brockley, T. (1984). The experience and effects of stress in teachers. *British Educational Research Journal*, 10, 83–87.
- Depolli, K. (1999). *Stres na delovnem mestu osnovnošolskega učitelja predmetnega pouka. [Stress in Primary Teachers]*. Neobjavljeno diplomsko delo, Filozofska fakulteta, Univerza v Ljubljani, Ljubljana, Slovenija [Unpublished bachelor’s thesis, Faculty of Arts, University of Ljubljana, Ljubljana, Slovenia].
- Dunham, J. (1992). *Stress in teaching*. London: Routledge.
- Esteve, J. (1990). Teacher burnout and teacher stress. V M. Cole in S. Walker (ur.), *Teaching and stress* (str. 4–25). Philadelphia: Open University Press.
- Galloway, D., Panckurst, F., Boswell, K., Boswell, C. in Green, K. (1987). Sources of stress for class teachers in New Zealand primary schools. *Pastoral Care in Education*, 5, 28–36.

- Jacobson, C., Pousette, A. in Thylefors, I. (2001). Managing stress and feelings of mastery among Swedish comprehensive school teacher. *Scandinavian Journal of Educational Research*, 45, 37–53.
- Jex, M. S. (1998). *Stress and job performance: Theory, research and implications for managerial practice*. Thousand Oaks, CA: Sage.
- Konrad, E. (1994). Osebnost v delovni situaciji [Personality in workplace]. V T. Lamovec (ur.), *Psihodiagnostika osebnosti 1* [Diagnostics of personality I], (str. 296 – 321). Ljubljana: Filozofska fakulteta, Znanstveni inštitut.
- Kyriacou, C. in Sutcliffe, J. (1977). Teacher stress: A review. *Educational Review*, 29, 299–306.
- Kyriacou, C. (1987). Teacher stress and burnout: An international review. *Educational Research*, 29, 146–152.
- Kyriacou, C. (1990). The nature and prevalence of teacher stress. V M. Cole in S. Walker (ur.), *Teaching and stress* (str. 27–35). Philadelphia: Open University Press.
- Kyriacou, C. in Sutcliffe, J. (1978). Teacher stress: Prevalence, sources, and symptoms. *British Journal for Educational Psychology*, 48, 159–167.
- Pečjak, V. (1977). *Psihologija spoznavanja* [Psychology of cognition]. Ljubljana: Državna založba Slovenije.
- Rudow, B. (1999). Stress and burnout in the teaching profession. V R. Vandenberghe in A. M. Huberman (ur.). *Understanding and preventing teacher burnout: A sourcebook of international research and practice* (str. 38–59) Cambridge: University Press.
- Skok, C. in Mihalič, J. (1996). Stresogeni dejavniki učiteljevega dela. [Stress factors among teachers]. *Šolsko svetovalno delo*, 1(2), (str. 15–20).
- Slivar, B. (2003). *Dejavniki, strategija in učinki spoprijemanja učiteljev s stresom* [Factors, strategies and effects of coping with stress among teachers]. Neobjavljeno doktorsko delo, Filozofska fakulteta, Univerza v Ljubljani, Ljubljana, Slovenija [Unpublished doctoral dissertation, Faculty of Arts, University of Ljubljana, Ljubljana, Slovenia].
- Travers, J. C. in Cooper, L. C. (1996). *Teachers under pressure*. London, New York: Routledge.
- Travers, J. C. in Cooper, L. C. (1998). Increasing costs of occupational stress for teachers. V J. Dunham in V. Varma (ur.), *Stress in teachers – Past, present and future* (str. 57–76). London: Whurr Publishers.
- Trenberth, L. in Dewe, P. (2005). An exploration of the role of leisure in coping with work related stress using sequential tree analysis. *British Journal of Guidance & Counselling*, 33(1), 101–116.
- Trenberth, L. in Dewe, P. (2006). Understanding the experience of stressors: The use of sequential analysis for exploring the patterns between various work stressors and strain. *Work & Stress*, 20(3), 191–209.
- Wolfgang, A. (1995). Job stress, coping, and dissatisfaction in the health professions: a comparison of nurses and pharmacists. V R. Crandall in P. L. Perrewé (ur.), *Occupational stress: A handbook* (str. 193–204). London, Washington: Taylor & Francis.
- Wolverton, M. in Wolverton, M. L. (2002). The interconnections between job satisfaction and work-related stress in academic deans. V S. G. Gordon in M. Wolverton (ur.), *Toward wellness – prevention, coping, and stress* (str. 101–125). Greenwich, CT: IAP.

Prispelo/Received: 12.05.2008

Sprejeto/Accepted: 15.07.2008

Recenzija znanstvene monografije “Psihološki vidiki preventivnih dejavnosti v prometu”

Melita Puklek*
Univerza v Ljubljani, Oddelek za psihologijo, Ljubljana

Review of the book “Psychological aspects of prevention in traffic”

Melita Puklek
Univerity of Ljubljana, Department of Psychology, Ljubljana, Slovenia

Ključne besede: prometna psihologija, prometna varnost, preventiva, Euchires

Key words: traffic psychology, highway safety, prevention, Euchires

CC = 4090

Zabukovec, V., Žlender, B., Polič, M., Divjak, M. in Markl, M. (ur.). (2007). *Psihološki vidiki preventivnih dejavnosti v prometu*. Ljubljana : Znanstvenoraziskovalni inštitut Filozofske fakultete. (COBISS.SI-ID: 238112768)

Pri Znanstvenoraziskovalnem inštitutu Filozofske fakultete v Ljubljani je spomladi 2008 izšla znanstvena monografija »Psihološki vidiki preventivnih dejavnosti v prometu« avtorjev Vlaste Zabukovec (ur.), Bojana Žlendra, Marka Poliča, Marka Divjaka in Mateje Markl. V monografiji so predstavljeni različni psihološki vidiki prometne varnosti, poglobljena analiza in strategije medijskih kampanj, uporabnost teorij učenja pri razvijanju prometno varnega vedenja pri otrocih in mladostnikih ter empirično evalvacijska študija, v kateri so avtorji predstavili mednarodni projekt Euchires, ki ga je v Sloveniji izvajal Svet za preventivo in vzgojo v cestnem prometu Republike Slovenije ter je bil usmerjen v spodbujanje ustrezne uporabe otroških varnostnih sedežev med vožnjo v osebnih avtomobilih pri otrocih.

* Naslov/Address: Melita Puklek, Univerza v Ljubljani, Filozofska fakulteta, Oddelek za psihologijo, Aškerčeva 2, 1000 Ljubljana, Slovenija, e-mail: melita.puklek@ff.uni-lj.si

Avtorji so vsebine znanstvene monografije predstavili v petih smiselno zaokroženih delih: (1) Prometna varnost, (2) Človek – dejavnik prometne varnosti, (3) Medijske kampanje, (4) Izobraževanje otrok za prometno varno vedenje in (5) Projekt Euchires.

B. Žlender v prvem poglavju predstavlja prometno varnost v Sloveniji in naredi primerjalno analizo z državami Evropske unije. Prav tako se osredotoči na vzroke prometnih nesreč ter strukturo in dejavnike prometne varnosti.

M. Polič se v drugem poglavju osredotoči na sam pojem prometne varnosti ter podrobno analizira vlogo človeka kot prometnega udeleženca. Dragocen doprinos védenju o psiholoških dejavnikih, ki vplivajo na povzročitev prometne nesreče, predstavlja avtorjeva razlaga vzrokov in osebnostnih lastnosti, ki so glavni dejavniki tveganega vedenja za volanom. Ugotavlja, da so prometne nesreče (med drugim) tudi posledica človekove socialne patologije. Avtor tudi smiselno razloži kognitivne procese, ki sodelujejo pri otroškem razumevanju prometa in vedenju v njem. Za strokovnjake, ki sodelujejo v izobraževanju in preventivi prometne varnosti, so pomembni podatki o alkoholiziranem stanju in njegovem deležu pri nastanku prometnih nesreč ter sistemski ukrepi, ki jih v različnih državah izvajajo v funkciji preprečevanja nesreč nastalih zaradi uživanja alkohola in drog.

M. Divjak v tretjem poglavju ugotavlja vlogo medijskih kampanj pri spodbujanju prometne varnosti. Medijske kampanje evalvira z različnih vidikov ter opredeli temeljne faze v procesu načrtovanja in izvajanja medijskih dejavnosti, podrobno operacionalizira pomen izbire ciljne populacije, posebno pozornost posveti oblikovanju temeljnega sporočila kampanje in izbiri medijskih orodij za širjenje sporočila o prometni varnosti. V prispevku so oblikovane tudi smernice za pripravo medijske kampanje na osnovi podrobne analize metaanalitičnih študij o učinkih kampanj na področju prometne varnosti na vedenje v prometu. Avtor prepleta pogloblitve teoretske pristope o preventivni dejavnosti v prometu, raziskovalna spoznanja tako tujih kot domačih raziskovalcev ter na osnovi raziskovalnih spoznanj izvede aplikativne predloge strategij oblikovanja dobre medijske kampanje.

V. Zabukovec v četrtem poglavju predstavi osnovna izhodišča izobraževanja za prometno varno vedenje ter izkušnje drugih evropskih držav na tem področju. Na podlagi temeljnih psiholoških teorij učenja – vedenjske, socialno-kognitivne in kognitivne teorije – predstavi načrt možnih intervencij v izobraževanju o prometni varnosti, zlasti otrok in mladostnikov, predvsem v smislu razvijanja znanja in spretnosti. Predstavljeni so tudi empirični rezultati študije EU ROSE 25, ki se v državah članicah Evropske unije ukvarja s kurikulum prometne vzgoje ter celovitim programom izobraževanja za prometno varno vedenje. Avtorica poudari temeljno vrednost projekta, ki je v tem, da mora izobraževanje za prometno varno vedenje biti del vseživljenjskega učenja ter vključevati tudi različne otrokove socialne partnerje, kot so starši, vrstniki, učitelji.

Najbolj raziskovalno usmerjeno poglavje je prispevek M. Markl, ki predstavlja projekt Euchires, katerega temeljni namen je ozaveščati otroke in starše o pomenu

otroških varnostnih sedežev in varnostnih pasov, informirati o pravilni uporabi le-teh in tako povečati varnost otrok med vožnjo. Avtorica poda znanstveno utemeljeno empirično analizo učinka medijske kampanje ter projektnega dela v vrtcih ter učinke izpeljanih delavnic na znanje, stališča in vedenje pri vzgojiteljicah in starših otrok. Spodbudni so rezultati raziskave, saj so pokazali, da so bile izpeljane dejavnosti učinkovite. Uporaba otroških varnostnih pasov pri drugem opazovanju se je povečala kar za 7,9 %.

Monografija Psihološki vidiki preventivnih dejavnosti v prometu predstavlja bogato in kompleksno zasnovano knjigo, ki za slovenski prostor ni razpoznavna le po njeni poglobljeni teoretski naravnosti, temveč predstavlja tudi originalne študije, ki so jih na področju preventivnih dejavnosti prometno varnega vedenja izvedli slovenski raziskovalci. Prepričana sem, da bo knjiga zanimiva tako za študente psihologije, študente pedagoških smeri kot tudi za različne strokovnjake, ki se v svojem delu ukvarjajo z informiranjem in ozaveščanjem o varnem vedenju v prometu ter preprečevanjem tragičnih posledic, ki jih ima nepremišljeno ravnanje v prometu.



zavod za fair play in strpnost v športu

SPORTIKUS



športna morala

Maja Smrdu
Stanislav Pinter
Milan Hosta



Ocena knjige “Skupinska psihoterapija: Od mitologije do teorije”

*Polona Matjan Štuhec**
Univerza v Ljubljani, Oddelek za psihologijo, Ljubljana

Review of the book “Group psychotherapy: from mythology to theory”

Polona Matjan Štuhec
University of Ljubljana, Department of Psychology, Ljubljana, Slovenia

Ključne besede: skupinska psihoterapija, skupinska dinamika, skupinsko terapevtski fenomeni, vloga terapevta

Key words: group psychotherapy, group dynamics, group therapeutic phenomena, therapist role

CC = 3313

Praper, P. (2008). Skupinska psihoterapija. Od mitologije do teorije. Ljubljana: Univerza v Ljubljani, Filozofska fakulteta. (ISBN: 978-961-237-192-0)

Vsi, ki kakorkoli klinično in svetovalno delamo s klienti v skupinah, z veseljem pozdravljamo knjigo z naslovom »Skupinska psihoterapija: Od mitologije do teorije« avtorja prof. dr. Petra Praperja, spec. klin. psih., psihoterapevta in skupinskega analitika. Pisec knjige je profesor klinične psihologije na Oddelku za psihologijo Filozofske fakultete v Ljubljani, poleg tega pa psihoterapevt z dolgoletno klinično prakso. S skupinami se je klinično pričel ukvarjati že v daljnem letu 1976, svoje skupinske psihoterapevtske izkušnje pa je dopolnil tudi s sistematično edukacijo iz skupinske analize Londonskega inštituta za skupinsko analizo. Vrsto let predava tudi izbirni predmet študentom psihologije z naslovom »Osnove skupinske psihoterapije«.

* Naslov/Address: Polona Matjan Štuhec, Univerza v Ljubljani, Filozofska fakulteta, Oddelek za psihologijo, Aškerčeva 2, Ljubljana, e-mail: polona.matjan@ff.uni-lj.si

Tako je knjiga »Skupinska psihoterapija« rezultat in nadgradnja avtorjevega lastnega raziskovalnega dela, poznavanja teoretičnih konceptov in spoznanj na področju delovanja terapevtskih skupin ter bogatih izkušenj iz njegove klinične prakse. S to knjigo smo dobili pregledno in obsežno delo na področju svetovalnega in terapevtskega dela s skupinami.

Knjiga je razdeljena v štiri dele. V prvem delu avtor bralca uvede v odkrivanje pomena in učinkov skupinske psihoterapije ter ga pouči o začetkih znanstveno utemeljene skupinske psihoterapije ter mu odkrije polpreteklo zgodovino na tem področju v Sloveniji. Skupinsko psihoterapijo in pristope klasificira po različnih kriterijih.

V drugem delu pisec predstavi posamezne teoretične koncepte in njihovo aplikacijo v skupini. Pregleda osnovne paradigme skupinske psihoterapije. Na prvo mesto postavi psihoanalitično paradigmo ter razdela modele, ki izhajajo iz nje in obsegajo psihoanalizo v skupini, analitično skupinsko psihoterapijo ter skupinsko analizo. Ker je skupinska analiza model, ki smo ga posvojili psihoterapevti v Sloveniji, ga avtor podrobneje predstavi.

Loti se pojma skupinske dinamike, ki je prav tako v Sloveniji prisoten že desetletja, pretežno v strokah, ki se ukvarjajo z oblikami psihosocialne pomoči, kot npr. v socialnem delu, domski pedagogiki, socialni pedagogiki. Opredeli intrapsihično dinamiko ter skupinsko dinamiko in pokaže, kako se ti dve prepletata. Nadalje Praper izpelje pristope, ki izhajajo iz skupinsko dinamičnih pristopov. Tako bralcu opiše osnove gestalt terapije in njeno aplikacijo v skupini, psihodramo kot tipično skupinsko modaliteto in skupinsko transakcijsko analizo. Ob vsaki vrsti skupinske psihoterapije najprej opiše njene teoretične osnove, razloži, kako posamezni teoretični pristop razume normalno osebnost in kako motenost. Nato se avtor loti pojasnjevanja terapije ter še posebej terapevtskih značilnosti v skupini.

Naslednje poglavje avtor posveti spoznanjem ego psihologije ter razvojne analitične psihologije. Že zgodaj so spoznanja ego psihologije vplivala na klinično prakso in brez spoznanj razvojne analitične psihologije je nemogoče v psihoterapijo sprejemati osebe z mejno organizacijo osebnosti.

Kot temeljna mehanizma v strukturalizaciji objektivnih odnosov avtor opiše mehanizma ponotranjanja in pozunanjanja, ki ju slikovito ilustrira z lastnim kliničnim primerom. Pokaže, kako se skozi teorijo objektivnih odnosov strukturira in razvije skupina ter opiše njene klinične aplikacije.

Pomemben prispevek k razumevanju osebnosti je dala self psihologija. Avtor jo opiše v posebnem poglavju, razloži pojme selfobjekt, self in selfobjektni odnos. Objekt je selfobjekt, kadar ga intrapsihično doživljamo tako, da nas oskrbuje s funkcijami, ki nam v odnosu vzbujajo, vzdržujejo in s pozitivnimi čustvi podpirajo občutenje sebe. Self psihologija je pomagala razumeti zaplete v ojdipski situaciji in doprinesla k razumevanju dinamike. Pisec razloži, kako psihologija selfa razume patologijo in kako poteka terapija.

Zadnji podpoglavji v tem delu sta namenjeni teorijam učenja in vedenjski

skupinski terapiji. Avtor razdeli vedenjske tehnike na tiste v nestrukturirani skupini in jih loči od onih v strukturirani skupini. Nadalje predstavi pisec humanistično paradigmo in z njo povezano skupinsko psihoterapijo ter eksistencialno paradigmo in skupinsko analizo. Knjiga se sistematično loteva vseh relevantnih konceptov, na koncu najdemo še teorijo socialnih sistemov ter kognitivni pristop, oboje kot izhodišča za skupinski tretma.

Tretji del knjige je posvečen skupinsko terapevtskim fenomenom. V knjigi so predstavljeni univerzalni skupinski fenomeni. Najbolj tipična skupinska fenomena, kljub temu da izhajata iz skupinske analize, sta postavljena na začetek, to sta skupina kot celota ter skupinski matriks. Celota v skupini se pokaže in vidimo jo kot komunikacijsko mrežo odnosov. Temeljni matriks je sestavljen prav iz komunikacije in interakcij v skupini, dinamični matriks pa se kaže skozi prizadevanje posameznika, da ta najde svoje mesto v skupini. Praper se dotakne velikosti skupin in kako le-ta vpliva na procese v njej ter kje lahko različno velike skupine koristno terapevtsko uporabljamo. Obdela načine komuniciranja ter se posebej posveti neverbalnim znakom in značilnostim neverbalne komunikacije. Spoznamo, kaj so zaprte in kaj odprte skupine in se srečamo s pojmom kohezivnosti in koherence. Ugotovimo, kako pomembno je zrcaljenje v skupini in kakšni patološki procesi so v ozadju malignega zrcaljenja.

V zadnjem delu se avtor loti vloge terapevta, razmišlja o tem, ali je ta univerzalna v različnih terapevtskih pristopih? Kako se terapevtova vloga spreminja skozi terapevtski proces v skupini? Tako kot je pisec prej opisal značilnosti skupin v različnih teoretičnih paradigmah, tako zdaj v luči istih paradigem pogleda terapevtovo osebnost, njegovo vlogo, način dela, skratka, dotakne se terapevtovih intervencij.

Knjiga zapolnjuje vrzel na področju dela s skupinami in bo zanesljivo pristala v rokah mnogih strokovnjakov, različnih skupinskih psihoterapevtov, pa tudi vseh drugih strokovnjakov, ki se ukvarjajo s skupinami – psihologov, psihiatrov, socialnih delavcev, socialnih pedagogov in drugih, ki se poklicno ukvarjajo s skupinami. Ker je knjiga pisana sistematično, tematika organizirana v smiselna poglavja, poglavja pa imajo jasne definicije in smiselne primere, bo gotovo postala koristen učbenik za vse študente in specializante, ki se usposabljujejo za terapevtsko delo s skupinami

V naslednji številki (letnik 17, številka 4, 2008) PSIHOLŠKIH OBZORIJ bodo predvidoma objavljeni naslednji prispevki [Forthcoming articles]:

Andrej KASTRIN

Odkrivanje zakonitosti in podatkovno rudarjenje v psihologiji: uporaba odločitvenih dreves za napovedovanje dosežka na Lestvici iskanja dražljajev

Darko LONČARIĆ in Cirila PEKLAJ

Proaktivna in defenzivna samoregulacija pri učenju

Svetlana KOLIĆ-VEHOVEC, Sonja PEČJAK in Neža AJDIŠEK

Razlike med spoloma v (meta)kognitivnih in motivacijsko emocionalnih dejavnikih bralnega razumevanja

Martina HORVAT

Poročilo z 8. psihološke konference Alpe-Jadran



Univerza v Ljubljani
FILOZOFSKA
FAKULTETA

RAZPRAVE
Filozofske fakultete

Maja Zupančič in Tina Kavčič

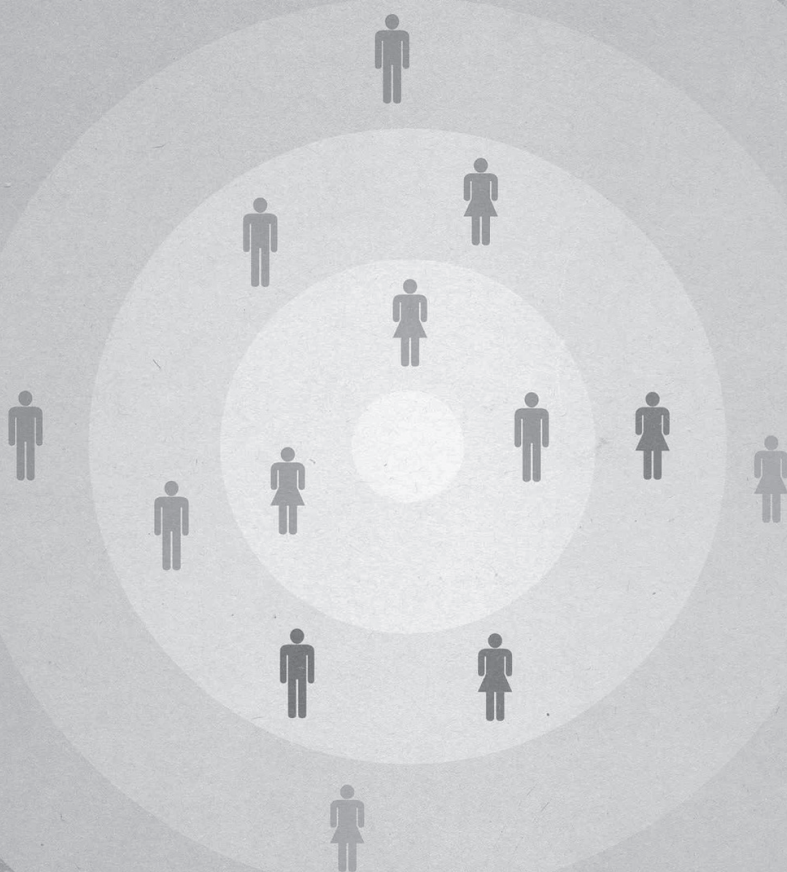


**Otroci od vrtca do šole:
razvoj osebnosti in
socialnega vedenja ter učna
uspešnost prvošolcev**

Ljubljana
2007

Valentina Hlebec,
Tina Kogovšek

MERJENJE SOCIALNIH OMREŽIJ



scripta

Navodilo avtorjem prispevkov

Opredelelitev prispevkov

V Psiholoških zbornikih so objavljeni prispevki, napisani v slovenskem ali angleškem jeziku (izjemoma, po presoji uredniškega sveta, še v kakšnem drugem svetovnem jeziku). Znanstveni in strokovni prispevki morajo biti napisani v jednatem, razumljivem, jasnem in logičnem ter jezikovno ustreznem slogu. Avtorji morajo sami poskrbeti za jezikovno lekturo prispevkov. *Empirični članek* avtor napiše kot poročilo o raziskavi. Vsebovati mora vse značilne dele, ki odsklikujejo korake raziskovalnega procesa: *uvod, metodo* (s podpoglavji *udeleženci, pripomočki oziroma instrumenti ter postopek*), *rezultate, razpravo in literaturo*. Za *teoretski članek* se predvideva, da v podrobnostih preuči in kritično analizira določene modele ali teorije. Empirični podatki so predstavljeni le, če so v neposredni zvezi s teorijo. *Pregledni članek* kritično ovrednoti na različnih mestih že objavljene prispevke. Značilni deli preglednega članka so opredelitev in razlaga problema, povzetek predhodnih raziskav, pojasnitev medsebojnih odnosov, protislovij, pomanjkljivosti in na koncu predlogi za nadaljnje raziskovanje.

Struktura prispevka

Prispevek mora vsebovati v slovenskem in angleškem jeziku napisane: *naslov* prispevka, *ključne besede* (okrog pet), ki najbolje opredeljujejo vsebino, in *povzetek* prispevka, ki vsebuje bistvene informacije o prispevku. Povzetek naj obsega do 300 besed (če je članek po dogovoru napisan v katerem od drugih svetovnih jezikov, morajo biti navedene informacije v tem, a tudi nujno v slovenskem in angleškem jeziku). V besedilu naj (zaradi postopka slepe recenzije) ne bo navedeno ime in/ali naslov avtorjev. Empirični prispevek naj v nadaljevanju sledi naslednjim splošnim pravilom:

1. Uvodni del ne sme biti preobširen, vendar mora ponuditi glavno teoretično ali konceptualno ogrodje, v katerega je vpet raziskovalni problem.
2. Raziskovalni problem mora biti jasno formuliran, predstavljati mora logično nadaljevanje uvodnega dela prispevka. Problem mora biti relevanten in v zastavljeni obliki še neraziskan. Poudarjeno mora biti, katere nove dileme odpira.
3. Raziskovalne hipoteze in spremenljivke morajo biti operacionalizirane in postopki opazovanja in merjenja morajo biti natančno opisani. Kratek, a natančen opis izbire udeležencev, uporabljenih psiholoških inštrumentov in raziskovalnega načrta je nujen.
4. Uporabo statističnih postopkov in rezultate morajo avtorji opisati s potrebno natančnostjo. Tabele morajo vsebovati le nujne informacije. Če je potrebno, so podrobnosti lahko predstavljene v dodatku (prilogi).
5. Razprava in interpretacija izsledkov se mora nanašati na znane koncepte in teorije, ne glede na to, ali jih dobljeni rezultati podpirajo ali ne.
6. Trditve ali dognanja drugih avtorjev so v besedilu potrjena z referenco. Na koncu prispevka je priložen seznam literature, na katero se besedilo sklicuje.

Za druge vrste sestavkov se ta struktura ustrezno in smiselno prilagodi.

Tehnična navodila za predložitev prispevkov

APA standardi

Obrlika poslanih rokopisov naj sledi standardom in priporočilom, opisanim v priročniku za pisanje raziskovalnih poročil, ki ga je izdalo Ameriško združenje psihologov (APA Publication Manual, 5. izd., 2001). Vsak rokopis, ki bo prispel v objavo v revijo Psihološka zbornika, ki bo ustrezal tehničnim (oblikovnim) zahtevam in bo sodil v okvir "namena in ciljev revije", bo po postopku dvojne slepe recenzije poslan v oceno dvema kompetentnima ocenjevalcema.

Oblikovni izgled

Avtorji besedilo prispevka pripravijo v enem od standardnih računalniških programov za obdelavo besedil (npr. Word za Windows) in ga shranijo v datoteki standardnega formata, npr. .doc ali .rtf. Besedilo naj bo napisano z dvovrstičnim ali 1,5 vrstičnim razmikom, z različnimi naslovi, podnaslovi in oštevilčenimi stranmi. Napisano naj bo z eno osnovnih oblik pisave (Times Roman, Arial, Helvetica, Courier) velikosti 12, besedilo naj bo levo poravnano. Začetki odstavkov naj ne bodo umaknjeni navznoter, pač pa naj bo pred vsakim novim odstavkom, naslovom, podnaslovom, predvidenim mestom za sliko ali tabelo vrinjena prazna vrstica. Prispevek naj ne bo daljši od dveh avtorskih pol (32 strani po 32 vrstic z okoli 60 znaki v vrstici oziroma 60.000 znakov, vključno z razmiki).

Tabele, slike, opombe

Za vsako tabelo in sliko mora biti v prispevku nakazano približno mesto, kjer je predvidena (npr. 'vstaviti sliko 1' ali 'vstaviti tabelo 1'). Tabel in slik avtorji ne vstavljajo v besedilo, pač pa jih pripravijo v ločeni datoteki, v katero najprej vstavijo vse tabele in nato vse slike, vsak prikaz na svoji strani (če je potrebno, lahko slike shranijo tudi v več ločenih datotekah).

Zaporedna številka slike in besedilo, ki sliko opisuje, naj bosta v besedilo vstavljena takoj za mestom, kjer je označen predvideni položaj slike (npr. najprej v svoji vrstici podamo informacijo: 'vstaviti sliko 1', pod njo pa ime in za piko naslov slike: 'Slika 1. Odnos med X in Y ...'). Slike morajo biti izdelane brezhibno in z dovolj velikimi črkami, številkami in ostalimi znaki (nabora Arial ali podobno), ki omogočajo pomanjševanje brez večje izgube preglednosti. Avtorji ne uporabljajo barv, pač pa, če je potrebno, različne dobro razločljive raste, sivine in/ali vzorce. Avtorji naj grafe v datoteko s tabelaričnimi in grafičnimi prikazi prilepijo v taki obliki, da jih bo možno urejati v izvornem programu (npr. Excelov grafikon naj prilepijo kot predmet in ne kot sliko ali metadatoteko). Fotografije z visoko ločljivostjo (> 300 dpi) naj bodo shranjene v kar najboljši kvaliteti grafičnega formata JPG, TIF ali PNG. Sheme in diagrami naj bodo iz izvornega programa prilepljene v datoteko kot predmet, izjemoma kot slike navedenih grafičnih formatov z visoko ločljivostjo. V primeru posebnih zahtev (npr. pri uporabi manj običajnih programov za generiranje slik ali če se določene slike ne da izdelati računalniško), naj se avtorji o načinu priprave slik predhodno posvetujejo s tehničnim urednikom revije (e-naslov: luka.komidar@ff.uni-lj.si).

Tabele naj bodo natipkane z enojnim razmikom. Nad tabelo naj bo (v datoteki s tabelaričnimi in slikovnimi prikazi) izpisana zaporedna številka tabele in za piko njen naslov (npr. Tabela 1. *Korelacije med ...*). Tabela mora biti informativna brez posebnega sklicevanja na

besedilo, torej opremljena s potrebnimi informacijami in po potrebi z opombami. Avtorji se v besedilu sklicujejo na sliko ali tabelo (npr. z "glej sliko 1", "v tabeli 2" ...), saj prikaz ne bo nujno na mestu, ki ga je predvidel avtor.

V poglavju o rezultatih naj bo isti podatek vedno prikazan le enkrat. Avtorji naj se odločijo, kateri način (slika, tabela ali prikaz v vezanem besedilu) je najbolj primeren in informativen. Opombe pod črto naj bodo vključene v prispevek le izjemoma.

Citiranje, literatura

Uporabljene reference drugih avtorjev naj bodo v besedilu citirane po harvardskem sistemu: npr. Rostohar (1952) ali (Rostohar, 1952). Kadar je citiranih več avtorjev, so navedeni v abecednem redu, npr. (Bujas, 1953; Rostohar, 1952; Trstenjak, 1953). Citati posameznih referenc so ločeni s podpičjem, npr. (Petrič, 1970; Petrovič, 1969). Kadar sta citirano delo napisala dva avtorja, se ves čas v prispevku navaja priimeka obeh avtorjev (npr. Schutz in Gessaroli, 1993). Kadar so citirano delo napisali trije, štirje ali pet avtorjev, so pri prvem citiranju vedno navedena imena vseh soavtorjev, npr. (Toličič, Šebek, Pečjak in Zorman, 1957), pri morebitnih naslednjih citatih pa le ime prvega avtorja, za druge pa je dodano le "idr."; drugi citat bi se tako glasil (Toličič idr., 1957). Kadar je citirano delo napisalo šest avtorjev ali več, se v vsem prispevku navaja le ime prvega avtorja in doda "idr.". Kadar je citiranih več del istega avtorja, napisanih v istem letu, so letnicam dodane male črke po abecednem redu, npr. (Peršič, 1968a, 1968b).

V seznamu literature na koncu prispevka so navedena po abecednem redu avtorjev (in brez zaporednih števil) vsa v besedilu citirana dela (in samo ta). Celoten seznam literature mora biti napisan v skladu z APA standardi citiranja. Pri navedbi vira so vedno izpisana imena vseh avtorjev prispevka, ne glede na to, koliko jih je.

Navedki prispevkov v revijah morajo vsebovati priimek avtorjev, začetnice imena, leto izdaje, naslov prispevka, polno (neokrajšano) ime revije (v poševnem tisku), letnik (v poševnem tisku), če se v vsakem zvezku znotraj istega letnika število strani začne z 1, tudi številko zvezka (v oklepaju, stičnim z letnikom, v navadnem tisku), in navedbo strani, na katerih je natisnjen prispevek (pri tem uporabljamo pomišljaj –, ne vezaj –). Paziti je potrebno na ločila, ki ločijo posamezne enote navedka. Primer navedbe:

Plomin, R. in Caspi, A. (1998). DNA and personality. *European Journal of Personality*, 12, 387–407.

Navedba avtorske knjige vsebuje priimek avtorjev, začetnice imena, leto izdaje, naslov knjige (v poševnem tisku), kraj izdaje in založbo. Primer navedbe:

Lazarus, R. S. (1991). *Emotion and adaptation*. Oxford: Oxford University Press.

Navedba poglavja avtorja v knjigi z urednikom vsebuje priimek avtorjev, začetnice imena, leto izdaje, naslov poglavja v knjigi, začetnice imena ter priimek urednikov, označbo, da gre za urednike, naslov knjige (v poševnem tisku), strani, na katerih je natisnjeno poglavje, kraj izdaje in založbo. Primer navedbe:

Schutz, R. W. in Gessaroli, M. E. (1993). Use, misuse and disuse of psychometrics in sport psychology research. V R. N. Singer, M. Murphey in L. K. Tennant (ur.), *Handbook of research in sport psychology* (str. 901–917). New York: Macmillan.

Vsaka navedba prispevka, katerega naslov ni v angleščini, mora imeti v seznamu referenc v oglatem oklepaju (v enakem tisku kot naslov prispevka) dodan tudi angleški prevod naslova prispevka. Primera navedbe:

Pogačnik, V. (1995). *Pojmovanje inteligentnosti [Conceptions of intelligence]*. Radovljica: Didakta.
Tušak, M. (1998). Barvne preference, simbolika barv in osebnost [Colour preferences, colour symbolism and personality]. *Psihološka obzorja*, 7(4), 67–79.

Oddajanje prispevkov

Avtorji besedilo in druge dele prispevka pripravijo v elektronski obliki. Datoteke, ki naj bodo poimenovane s priimkom prvega avtorja in dodano specifično oznako (npr. novak-besedilo.doc; novak-tabeleslike.doc), avtorji pošljejo glavni in odgovorni urednici na elektronski naslov anja.podlesek@ff.uni-lj.si.

Ob prvem pošiljanju prispevka avtorji pripravijo dodatno datoteko (npr. novak-kontaktnipodatki.doc) z osnovnimi podatki o prispevku in avtorjih: izpišejo naslov prispevka, ime in priimek avtorjev, strokovne nazive, ime inštitucije, v kateri so zaposleni (v slovenščini in angleščini), in natančen naslov tistega avtorja, s katerim bo uredništvo revije komuniciralo (tudi elektronski naslov, številko telefona in, če je možno, telefaksa ter URL naslov).

Če bo potrebno (npr. v primeru, da prispevek vsebuje veliko simbolov, ki jih računalniki recenzentov ne bi ustrezno interpretirali, ali v primeru, da recenzenti želijo pregledovati natisnjeno obliko prispevka), bo urednica naknadno zпросila še za natisnjeno obliko prispevka.

Po končanem redakcijskem postopku in strokovnih recenzijah bo avtor prejel recenziji prispevka in kratko mnenje urednice glede sprejetja besedila v objavo. V primeru, da je prispevek sprejet v objavo, avtorji upoštevajo vse prejete pripombe, popravke in sugestije ter pripravijo končno verzijo prispevka. V končni verziji naj bo prva stran besedila takoj za naslovom prispevka dopolnjena še z imenom in priimkom avtorjev, imenom in krajem inštitucije, kontaktnimi podatki in morebitnimi dodatnimi informacijami o financerju studije, o tem, da je bil prispevek predstavljen na kakšnem od kongresov, ali zahvalo. Za angleškim naslovom naj bodo dodani imena in priimki avtorjev ter ime, kraj in država njihovih institucij v angleščini.

Avtorji tudi končno verzijo prispevka oddajo glavni in odgovorni urednici v elektronski obliki. Če je potrebno, natisnejo en izvod prispevka in ga pošljejo na naslov urednice (v tem primeru bodo tipkopis in slike mesec dni po objavi uničeni, če avtorji ne bodo posebej pisno zahtevali vrnitve originalnih gradiv).

Zaključne opombe

Poslana končna verzija rokopisa pomeni tudi potrditev avtorjev, da prispevek v enaki ali podobni obliki ni bil objavljen v kateri drugi domači ali tuji publikaciji in da tudi v bodoče ne bo brez poprejšnjega soglasja izdajatelja Psiholoških obzorij. S spletnih strani <http://psy.ff.uni-lj.si/i/Guests/Obzorja/Avtorjem/avtorjem.html> avtorji natisnejo, izpolnijo in podpišejo Obrazec za odstop avtorskih pravic ter ga po navadni pošti pošljejo urednici. Uredniški odbor, uredniški svet ter izdajatelj ne prevzemata odgovornosti za strokovna mnenja in trditve oziroma zaključke, ki so jih podali avtorji v posameznih prispevkih.

Instructions for Authors

Papers should be written either in Slovenian or English language (occasionally, the Scientific Board might also accept for publication papers in other languages). Scientific and technical papers should be written in economic, intelligible, clear and concise style. An *empirical paper* should report original research, presenting all the standard elements of scientific investigation (introduction, method — including *participants, instruments and procedure* - results, discussion, references). A *theoretic paper* is expected to examine in detail and critically analyse selected models and/or theories, and empirical data are described only if they are directly related to the theory. A *review paper* is expected to evaluate previously published work and it is typically composed of the following sections: problem definition, summary of previous research, explanation of subject matter inter-relations, contradictions, problems, and suggestions for further research. *Meta-analytic study* is a particular type of article, based on the established meta-analysis methodology, comparing different empirical investigations addressing a common problem.

The manuscript must contain the English and Slovenian version of the *title, keywords* (about five, defining the contents) and *abstract* in up to 300 words (the authors that do not speak Slovene should ask the Editor for help). If the paper is written in some other world language, the title, keywords, and abstract must be provided in that language in addition to the English and Slovenian version.

A typical empirical paper should be written in accordance with the following guidelines:

1. Introduction should not be too extensive, yet it should provide the necessary conceptual framework.
2. The problem should be clearly and consistently defined, following logically from the introduction. It should be of sufficient relevance and novelty, whereby the new dilemmas opened should be emphasised.
3. Research hypotheses and variables of interest should be concisely defined; the observation and measurement procedures should be precisely described. A short but accurate description of psychological instruments applied and of the methodology in general (research design, selection of participants) is mandatory.
4. Application of statistical techniques should be described in sufficient detail. The tables and figures may only contain essential information (if necessary, details can be presented separately in appendix).
5. Discussion and interpretation of the findings should refer to the established concepts and theories, regardless of whether they are supported by the findings or not.

The Editorial and Scientific Board assume no responsibility for the expert opinions, claims, and conclusions stated by the authors in their papers.

Technical Aspects of Manuscript Preparation

The manuscripts should be prepared in accordance with the *APA Publication Manual* (Fifth Edition, 2001). Each manuscript, meeting the technical standards and falling within the aims and scope of the journal, will be subject to double-blind review by two reviewers.

The manuscript must be typeset in a standard font (Times Roman, Arial, Helvetica, or Courier, 12 pt size), with left justification and double spacing. Titles and subtitles must be clearly indicated and pages must be numbered (no other header/footer information is allowed). The paragraphs must not be indented and they must be separated by a blank line. The paper should normally not exceed 32 standard pages (32 rows by 60 characters) in length. The main text, without author(s) name(s) and institution(s), should be saved in a .doc or .rtf file (e.g., in MS Word – if non-Windows software is used, the author must perform the necessary conversion to assure seamless operation on the Windows platform). Tables and figures should not be included in the text (but indicate their position). They should be added as separate files in a standard graphics format (EPS, JPG, PDF, PS, BMP, GIF, TIF etc.). Another separate file must be added containing the basic information on the paper – name(s) of the author(s), title(s), institution(s), title of the paper and exact address of the author the Editorial Board will contact for further correspondence (including telephone number, fax number, e-mail address, home page URL if possible etc.). The files should be named by the first author's surname and a short description of the file (e.g. socan-text.doc, socan-figure1.jpg, socan-about_authors.doc). All the files should be sent by e-mail to anja.podlesek@ff.uni-lj.si

On Editor's request, the printed version (three copies) of the manuscript should be sent to: Dr. Anja Podlesek, University of Ljubljana, Faculty of Arts, Department of Psychology, p. p. 580, SI-1001 Ljubljana, Slovenia. Each table and/or figure with the proper caption should be provided on a separate sheet. Provide figures in camera-ready format without excessively small details. The manuscript will be destroyed one month after publication unless the author submits a request in writing for the originals to be returned to her/him.

After editorial consideration and review process a copy of the manuscript will be returned to the author together with the two reviews and editorial opinion on acceptability. In case the manuscript is accepted for publication, the author must take into account all the suggestions (including language and style changes) while preparing the final version. For all other aspects of the final version the instructions for the initial version apply, except that the authors' names and institutions should be added into the text and contact information should be added in a footnote. Additional information on funding and/or congress presentation and acknowledgements should be entered in a footnote, which will be attached to the paper's title.

Submission of the final version implies author's declaration that the paper has not been published in any other Slovenian or foreign publication in its present (or similar) form. At the same time it implies author's consent that the paper or its parts can only be reproduced in any other publication by permission of the publisher of the Horizons of Psychology. The 'Consent to Publish & Transfer of Copyright' form available at <http://psy.ff.uni-lj.si/Guests/Obzorja/Instructions/instructions.html> should be printed, signed and posted to Editor-in-Chief.

For any additional correspondence please make use of e-mail.