Intercultural Sensitivity Scale:
Proposal for a Modified Serbian Version

Danijela S. Petrović[1], Jelena Starčević[2], Guo-Ming Chen[3], and Darko Komnenić[1]

[1]*Department of Psychology, Faculty of Philosophy, University of Belgrade, Serbia*
[2]*Faculty of Pedagogical Sciences, University of Kragujevac, Serbia*
[3]*Department of Communication Studies, University of Rhode Island, USA*

Intercultural Sensitivity Scale (ISS) is the main assessment tool for measuring intercultural sensitivity as an affective component of intercultural communication competence. The ISS has been developed based on an American sample and therefore there is a need to check possibilities for its application in another cultural context. In this study, we tested whether the factor structure of the original scale is confirmed in a Serbian sample as well. The results show that the compatibility of factor structure is not satisfactory ($\chi^2/df = 3.38$; CFI = .78; RMSEA = .07) and the application of the scale requires modification. A proposal for a modified version of the ISS is presented together with evidence for its usage. The main advantages of the modified version are: (a) a corresponding factor structure, (b) higher internal consistency and (c) better prediction of relevant criteria.

Keywords: intercultural sensitivity, ISS, psychometric properties

There are various definitions of intercultural sensitivity in psychology. It is, for example, defined as an "ability to discriminate and experience relevant cultural differences" (Hammer, Bennet, & Wiseman, 2003, p. 422), while a certain cognitive structure or worldview configuration is considered to be its essence (Bennett & Bennett, 2004). Besides, intercultural sensitivity is determined as a complex psychological disposition which includes an interest in other cultures, sensitivity to notice cultural differences, and willingness for modification of behavior as an indication of respect for the people of other cultures (Bhawuk & Brislin, 1992). However, Chen and Starosta (2000a) made a call for conceptual clarity and argued for a more specific view of intercultural sensitivity. They acknowledged that intercultural sensitivity is related to three aspects of intercultural interaction (cognitive, affective and

behavioral), but they emphasized that it *mainly deals with affect*. These authors defined intercultural sensitivity as "an individual's ability to develop a positive emotion towards understanding and appreciating cultural differences in order to promote appropriate and effective behaviour in intercultural communication" (Chen & Starosta, 2000a, p. 408). Additionally, Chen and Starosta (2000a, 2000b) described six components of intercultural sensitivity: self-esteem, self-monitoring, open-mindedness, empathy, interaction involvement and suspending judgment. These components form the fundamental bases for positive emotional reactions toward people from other cultures in the process of interaction.

In order to assess intercultural sensitivity, Chen and Starosta (2000b) developed the Intercultural Sensitivity Scale (ISS). After preliminary analyses, they retained 44 out of 73 items in the pilot study. In order to determine the factor structure of the scale they conducted a new study on the sample consisting of 414 college students enrolled in basic communication courses (63% female and 37% male). The average age of the participants was 20.65. In the analysis of the collected data, Chen and Starosta used Principal Axis Factoring (PAF) with oblique rotation. Five factors with Eigenvalues of 1.00 or higher were extracted and these factors accounted for 37.3% of the variance. The items with factor loadings of at least .50 and with secondary loadings no higher than .30 remained in the final version of the scale. This version is comprised of 24 items. Seven items were included in the first factor. Most of these items are concerned with participant's feeling of participation in intercultural communication. This factor was labelled "Interaction Engagement". Six items were clustered in the second factor. These items mainly express the way participants orient to or tolerate their counterpart's culture and opinion. This factor was labelled "Respect for Cultural Differences". Five items had a significant loading on the third factor. These items are concerned with how confident participants are in the intercultural setting, so the factor was named "Interaction Confidence". Three items were significantly loaded on the fourth factor. These items deal with the participant's positive or negative reaction towards communicating with people from different cultures, so this factor was defined as "Interaction Enjoyment". Finally, three items were clustered in the fifth factor. The fifth factor items refer to the participant's effort to understand what is going on in an intercultural interaction. The factor was labelled as "Interaction Attentiveness".

The ISS has demonstrated a good internal consistency. Cronbach's alpha reliability coefficients were between .79 and .89 (Chen & Starosta, 2000b; Graf & Harland, 2005; Petrović & Zlatković, 2009). Cronbach's alpha coefficients for the four ISS subscales (subscales represent factors in the structure of the instrument) ranged from .70 do .75, while alpha coefficient of "Interaction Attentiveness" was .47 (Graf & Harland, 2005). A five-factor scale structure was initially confirmed on a German sample by the means of confirmatory factor analysis (Fritz, Möllenberg, & Chen, 2002). However, the findings of this study indicated minor weaknesses in the operationalization of the constructs underlying the instrument, e.g. the reliability of several items was not sufficiently high. In a later study with two matched samples, the one from USA ($n_1 = 188$) and the

other from Germany (n_2 = 179), the findings did not sufficiently verify the five-factor structure of ISS (Fritz, Graf, Hentze, & Möllenberg, 2005). At the same time, Fritz et al. (2005) admit some limitations of their study. The sample in their study differs from the samples in previous studies. The number of participants in the matched samples is half the number of participants in previous studies. This is important because sample size can influence the parameter estimates (see also Rusell, 2002). Further, some analyses were done on the overall sample which combines German and American participants and some were done on a reduced number of items. The reduction of number of items might negatively affect some indicators of model fit in terms of factor reliability and average variance. Apparently, more studies are needed to explore the structural validity of ISS, especially for participants from non-Western cultures.

The construct validity of the ISS (i.e. its convergent and discriminant validity) has been established through several studies on the basis of the height of correlation between the ISS and other scales – for example, "Self-Monitoring Scale", "Intercultural Effectiveness Scale", "Interpersonal Competence Questionnaire" (Chen & Starosta, 2000b; Graf & Harland, 2005). Graf and Harland (2005) found that the ISS scales, with the exception of "Interaction Confidence", were significantly correlated with intercultural decision quality in a problem scenario. In addition, a regression including five intercultural and interpersonal scales showed that the ISS total score was a statistically significant predictor of intercultural decision quality (β = .15, p <.05).

Both construct and predictive validity of the ISS were investigated in USA only. It is an open question to what extent one instrument evaluates the same psychological dimension when applied to a new lingua-cultural milieu. Greenholtz (2005) noted that some authors unduly rely on the coefficient of internal consistency as the sole criterion of the instrument`s adjustment for appliance in new cultural context. In order to make a decision on the use of the instrument, it is also necessary to collect the data on other psychometric characteristics, with the focus on various types of validity (its structural, construct and criterion validity).

The Present Study

The ISS is the only scale with a primary goal to measure the emotional aspects of intercultural competence. Moreover, there is no such instrument in the Serbian language with a purpose of assessment of any sort of intercultural competencies. The main purpose of this study is therefore to check and analyse the possibilities of having ISS applied in Serbia. In this regard, we examined the accordance between the factor structure of the ISS on a Serbian sample and the five-factor structure of the ISS developed by Chen and Starosta (i.e. the structural validity of the instrument). In case it proves that the ISS has an inadequate model fit on the Serbian sample, an additional aim was to modify the scale in order to identify the best possible fitting model. The final aim was to examine and report the distributional properties, internal consistency, construct and predictive validity

of the scale and possibly its modified version. The construct and predictive validity of the scale were analysed in relation to measures of cultural intelligence, a construct which also expresses individual differences within intercultural interaction. According to initial evidence, The Cultural Intelligence Scale has high reliability and it also has structural validity (Petrović & Komnenić, 2012; Starčević, 2013) and construct validity (Starčević, 2013) on Serbian samples. Score prediction on motivational component of the scale could be of particular importance because it is conceptually the most similar to the ISS. Motivational subscale assesses the capability to direct attention and energy toward functioning in situations characterized by cultural differences (Ang et al., 2007).

Method

Participants and Procedure

The participants were 522 students of social sciences at the University of Belgrade. Among them 375 were female (72%) and 147 were male (28%). The average age of participants was 23.14 (SD = 2.87). The majority of participants (96%) have a Serbian nationality. The data was collected in a group setting. Participation in this study was voluntary and participants remained anonymous.

Instruments

Details about the ISS developed by Chen and Starosta were already provided in the introductory section. The *back translation method* was employed before administering this 24-item intercultural sensitivity questionnaire as one of the usual ways of scale translation and adaptation (Greenholtz, 2005). The instrument was first translated into the target language and then translated back to the source language by an independent translator. By comparing original and back translated versions of the instrument, the subject matter experts revealed some translation problems and solved them in cooperation with the translators.

The Cultural Intelligence Scale (CQS) is a self-report scale designed to measure metacognitive, cognitive, motivational and behavioral dimensions of cultural intelligence. According to a recent review, CQS is one of the three most promising instruments for assessing cross-cultural competence (Matsumoto & Hwang, 2013). The CQS consists of 20 items given along with a 5-point Likert-type scale where the participants express their own degree of agreement. In this research, Cronbach's alpha for the total scale is .82 and for the subscales are as follows: .69 for metacognitive, .74 for cognitive, .77 for motivational, and .82 for behavioural.

Results

Exploratory Factor Analysis of ISS

Exploratory factor analysis of the collected data was the first step in the examination of the ISS factor structure. In accordance with the procedure applied by Chen and Starosta (2000b) we used the PAF to explore the factor structure. As in Chen and Starosta's study, the five factors with Eigenvalues over 1 were extracted. Very similar results were found with oblique and orthogonal

factor rotation, but what will be presented is a solution obtained with orthogonal (varimax) rotation which is easier to interpret. These factors accounted for 36% of the variance and they are listed in order presented in Table 1: "Interaction Enjoyment", "Interaction Engagement", "Respect for Cultural Differences", "Interaction Confidence", and "Interaction Attentiveness". The item loadings for the five factors in the rotated solution are shown in Table 1.

Table 1
Rotated Factor Matrix with Factor Loadings on Five Factors of the ISS

Items	Factor				
	1	2	3	4	5
12. I often get discouraged when I am with people from different cultures.	**.682**	.050	.113	.208	-.069
15. I often feel useless when interacting with people from different cultures.	**.616**	.071	.151	.162	.101
9. I get upset easily when interacting with people from different cultures.	**.542**	.097	.304	.228	-.081
22. I avoid those situations where I will have to deal with culturally-distinct persons.	**.480**	.290	.203	.106	-.137
24. I have a feeling of enjoyment towards differences between my culturally-distinct counterpart and me.	-.051	**.588**	.117	.209	-.002
21. I often give positive responses to my culturally different counterpart during our interaction.	.283	**.531**	.057	.099	.104
1. I enjoy interacting with people from different cultures.	.106	**.498**	.290	.255	.099
17. I try to obtain as much information as I can when interacting with people from different cultures.	.101	**.464**	.135	.050	.092
23. I often show my culturally-distinct counterpart my understanding through verbal or nonverbal cues.	.027	**.456**	.002	.039	.109
16. I respect the ways people from different cultures behave.	.056	**.399**	.355	-.043	-.091
20. I think my culture is better than other cultures.	.152	**.367**	.333	-.101	-.248
8. I respect the values of people from different cultures.	.104	.060	**.673**	.018	.083
7. I don't like to be with people from different cultures.	.398	.200	**.504**	.080	.154
13. I am open-minded to people from different cultures.	.211	.332	**.490**	.042	-.112
18. I would not accept the opinions of people from different cultures.	.238	.228	**.395**	-.022	-.314
2. I think people from other cultures are narrow-minded.	.084	.071	**.327**	.005	-.108
3. I am pretty sure of myself in interacting with people from different cultures.	.177	.154	.013	**.759**	.049
10. I feel confident when interacting with people from different cultures.	.192	.067	.021	**.696**	-.001
5. I always know what to say when interacting with people from different cultures.	.102	.112	-.044	**.590**	.103
4. I find it very hard to talk in front of people from different cultures.	.309	.051	.316	**.355**	-.146
14. I am very observant when interacting with people from different cultures.	-.073	.071	.041	-.043	**.549**
11. I tend to wait before forming an impression of culturally-distinct counterparts.	.045	.081	-.046	-.006	**.408**
19. I am sensitive to my culturally-distinct counterpart's subtle meanings during our interaction.	-.089	-.011	-.137	.051	**.324**
6. I can be as sociable as I want to be when interacting with people from different cultures.	.133	.077	.039	.169	**.270**

Table 2 compares the data obtained in Chen and Starosta's (2000b) study and current exploratory factor analysis of ISS – according to the items of the

five factors. Bolded items are the items with the highest factor loadings on the same factors identified by Chen and Starosta in their study. Seventeen items out of 24 (or 70.8% of the scale) had their largest loadings on factors as expected, but only 11 items satisfied Chen and Starosta's criterion – items having loadings of at least .50 with secondary loadings no higher than .30.

Table 2
List of Items by ISS Factors: Chen and Starosta's Study (2000b) and Current Study

The Arrangement of Items	Factor				
	Interaction Engagement	Respect for Cultural Differences	Interaction Confidence	Interaction Enjoyment	Interaction Attentiveness
Chen and Starosta's Study	1,11,13,21, 22,23,24	2,7,8,16, 18,20	3,4,5,6,10	9,12,15	14,17,19
Current Study	**1,21,23,24,** 16,17,20	**2,7,8,18,** 13	**3,4,5,10**	**9,12,15**, 22	**14,19,** 6,11

Confirmatory Factor Analysis of ISS

The ISS factor structure was further tested by means of the confirmatory factor analysis (CFA). The data was analysed by LISREL program (version 9). The same procedure was used by Fritz et al. (Fritz et al., 2005; Fritz et al., 2002) in testing generalizability of the ISS.

Table 3 shows parameter values of model fit obtained in the current study and values obtained in two earlier studies on the ISS structure (Fritz et al., 2005; Fritz et al., 2002). Table 3 also presents recommended values of the parameters which serve as guidelines for Fritz et al. to evaluate the model fit. Global adjustment refers to overall measures of model fit, while local fit deals with the measures of the fit of model's parts. Chi-square is divided by its degrees of freedom (known as relative/normed chi-square) in order to get a statistic that minimizes the impact of sample size. Underlined values fail to meet the requirements.

The results of the CFA presented in Table 3 show that only one index of global fit meets the requirement. The averaged values of parameters of local fit fail to meet the requirements to a large extent, also. For example, the averaged value of indicator reliability (.29) is significantly below the recommended value (.40). Besides, the chi-square value is statistically significant, $\chi^2 = 818.47$, $df = 242$ ($p < .001$), which is contrary to the expectations when the model fits the data.

Modified Version of the ISS. The obtained results, especially the considerable mismatch between the ISS model observed in our study and the expected model of the ISS (as given in Chen and Starosta' study), implicate that the use of this scale in another cultural context requires its modification. In the next step we tested the reduced and adjusted model of the ISS. First of all, we excluded the items with very low reliabilities (near zero): 6, 11, 14 and 19. Once we excluded the two items of the "Interaction Attentiveness" factor, we did the same with item 17, which was the only remaining item that comprised this factor in the original scale structure. We also excluded the items 4, 7, 16

and 20 for the reason of split factor loading, according to the results of the PAF and to the results of the CFA which also demonstrate their relation to the various factors (i.e. the CFA results suggest possible "paths" between items and factors). Compared to the original scale structure, we changed the position of items 13 and 22, following the recommendations from CFA. This procedure should result in decreased value of the chi-square statistic. The applied changes are also in accordance with the results of the PAF shown in Table 1. The new model has 15 items and four factors in the structure:

1. "Interaction Enjoyment" with items 9, 12, 15 and item 22 (added item).
2. "Interaction Engagement" with items 1, 21, 23 and 24.
3. "Respect for Cultural Differences" with items 2, 8, 18 and item 13 (added item).
4. "Interaction Confidence" with items 3, 5 and 10.

The values of fit indices for this version of the scale are presented in Table 3 together with the values of fit indices for the original version of the scale (both in this study and in the studies of Fritz et al.).

Table 3
Empirical Models Comparison

Measures of fit	Requirement	Current study Original scale	Modified scale	Fritz et al., 2002	Fritz et al., 2005
Global Fit					
χ^2/df	≥2.50	3.38	2.25	1.96	1.84
RMSEA	≥.05	.07	.05	.05	.04
GFI	≥.90	.87	.95	.92	.92
AGFI	≥.90	.84	.93	.90	.89
RMR	<.10	.07	.05	.04	.03
NFI	≥.90	.72	.89	-	.80
NNFI	≥.90	.75	.91	-	.88
RFI	≥.90	.68	.86	-	.77
IFI	≥.90	.79	.93	-	.90
CFI	≥.90	.78	.93	-	.90
AIC	min.	1042.58	260.86	-	474.67
ECVI	min.	2.00	0.50	-	.130
Local Fit			**(Average)**		
Indicator Reliability	≥.40	.29	.36	.	.34
Factor Reliability (CR)	≥.60	.59	.68	.69	.65
Average Variance Extracted (AVE)	≥.40	.30	.38	.36	.31
Convergent Validity if CR	≥.60	.59	.68	.69	.60
if AVE	≥.50	.30	.38	.36	.31
Discriminant Validity (Fornell/Larcker Criterion)	AVE > r^2	.30>.27	.38>.26	.36>.25	.31<.39

χ^2=chi square statistic; df=degrees of freedom; RMSEA=root-mean-square error of approximation; GFI=goodness of fit index; AGFI=adjusted goodness of fit index; RMR=root mean square residual; NFI=normed fit index; NNFI=non-normed fit index; RFI=relative fit index; IFI=incremental fit index; CFI=comparative fit index; AIC=Akaike information criterion; ECVI=expected cross-validation index; r^2=squared correlations between two factors

Table 3 indicates that the newly obtained fit indices of the modified scale generally meet the criteria. This is especially the case with parameters of global fit.

Table 3 also shows parameters of convergent and discriminant validity of the scale factors. The convergent validity of the scale factors, expressed through their reliability (CR), can be regarded as sufficiently high, which is not the case when assessed on the basis of their average variance extracted (AVE). Fritz et al. (Fritz et al., 2005; Fritz et al., 2002) found the same, where reliability of factors is considered to be more important (Fritz et al., 2002).

According to the Fornell and Larcker (1981) discriminant validity criterion – the average variance extracted in the composite of items, i.e. factor, has to be higher than the squared correlations between the factors – discriminant validity of the scale factors is sufficient.

The value of chi square decreased but remained statistically significant, χ^2 = 199.96, df = 84 (p <.001).

Comparison Between the Two Versions of ISS

Further on, a comparison was made between the original and modified versions of the ISS with regard to four psychometric properties: (a) normality of data distribution, (b) internal consistency, (c) construct validity, and (d) predictive validity.

Distributional Properties. The data distribution parameters for two versions of the scale as well as for the single factors in the scale structure are given in Table 4. The data distribution in all cases differs from the normal one and it is negatively skewed. However, the skew values are relatively low (<1), except for the Interaction Enjoyment and Interaction Enjoyment Modified.

Table 4
Distributional Properties and Reliability of ISS: Original and Modified Versions

	ISS	ISS Modified	Interaction Engagement / Interaction Engagement Modified	Respect for Cultural Differences / Respect for Cultural Differences Modified	Interaction Confidence / Interaction Confidence Modified	Interaction Enjoyment / Interaction Enjoyment Modified	Interaction Attentiveness
Skewness	−0.57	−0.54	−0.49 / −0.46	−0.90 / −0.98	−0.21 / −0.46	−1.69 / −1.61	−0.12
Kurtosis	.241	.005	.123 / .264	.744 / .722	-.408 / .100	3.595 / 3.198	.253
Kolmogorov-Smirnov	.07	.07	.09 / .10	.13 /.16	.08 /.12	.21 / .20	.11
Sig.	.000	.000	.000 / .000	.000 / .000	-.000 / .000	.000 / .000	.000
Reliability α	.79	.78	.62 / .65	.66 /.59	.66 / .75	.72 / .73	.21

Reliability. The values of the alpha coefficients for the two scale versions and for single factors in the scale structure are also shown in Table 4. Although the modified scale factors mainly include a smaller number of items, their alpha

coefficients are generally higher except for the case of "Respect for Cultural Differences".

Construct validity. The correlations between the ISS subscales and the CQS subscales are shown in Table 5 as well as the correlations between the ISS modified subscales and the CQS subscales. Although the ISS modified subscales in general include a smaller number of items, their correlations with the CQS subscales are predominantly higher or the same value with regard to the correlations of the original ISS subscales.

Table 5
The correlations between the subscales: ISS, ISS Modified and CQS

	Meta CQ	Cog CQ	Mot CQ	Bih CQ
Enjoyment / Enjoyment Modified	.19**/.24**	.22**/.24**	.33**/.38**	-.06/-.05
Engagement / Engagement Modified	.43**/.39**	.32**/.35**	.56**/.56**	.21**/.25**
Respect / Respect Modified	.20**/.18**	.07/.09*	.33**/.31**	.02/.04
Confidence / Confidence Modified	.33**/.30**	.40**/.41**	.37**/.36**	.06/.06
Attentiveness	.21**	.14**	.11**	.22**

Predictive validity. The advantage of the modified version of the ISS is ultimately tested by checking its predictive validity for the global CQS score and for the scores on the CQS subscales. The predictors in the regression equations represent the sum of raw scores on the single factors in the ISS structure. In all hierarchical regression models the first block of predictors consists of the original scale factors and the second block of predictors consists of both original and modified scale factors.

Table 6
The CQS Score Prediction: Global Score and Individual Subscales

Dependent Variable	Model	R	R^2	Adjusted R^2	R^2 Change	F Change	$df1$	$df2$	Sig. F Change
CQS	Level 1	.606	.368	.361	.368	59.980	5	516	.000
	Level 2	.646	.417	.407	.049	10.821	4	512	.000
metaCQS	Level 1	.479	.229	.222	.229	30.718	5	516	.000
	Level 2	.485	.235	.222	.006	.980	4	512	.418
cogCQS	Level 1	.458	.210	.202	.210	27.393	5	516	.000
	Level 2	.496	.246	.233	.036	6.124	4	512	.000
motCQS	Level 1	.594	.353	.347	.353	56.311	5	516	.000
	Level 2	.633	.400	.390	.047	10.127	4	512	.000
bihCQS	Level 1	.298	.089	.080	.089	10.038	5	516	.000
	Level 2	.333	.111	.095	.022	3.189	4	512	.013

Level 1: original ISS factors
Level 2: original and modified ISS factors

The results show that the increment of the explained variance obtained on the second level of hierarchical regression is statistically significant in

almost every regression model, except when the criterion is metacognitive CQS subscale. The modified version of the ISS has the best results in predicting global CQS score and motivational CQS score.

Discussion

The ISS is an instrument developed in the USA for the assessment of the affective dimension of intercultural competencies (Chen & Starosta, 2000b). In this paper we analysed whether the ISS would be applicable in Serbia. The factor structure of the scale was analyzed in detail to enquire whether it corresponds to the structure originally defined by Chen and Starosta (2000b) when developing the instrument. In their review of the cross-cultural instruments, Matsumoto and Hwang (2013) indicated that the evidence for the structural validity of a number of these instruments is lacking, including for the ISS.

With exploratory factor analysis (PAF) we extracted five factors, as expected, which could be interpreted and labelled as in the original study. These five factors accounted for a rather small percentage of the variance, but similar to that reported by Chen and Starosta (2000b). The results, therefore, indicate that ISS is a heterogeneous instrument. However, the factors were not saturated by identical ISS scale items according to the structure its authors reported, and the factor structure of the scale was not simple, due to the fact that several items saturated more than one factor.

The concordance of the factor structure of the ISS as originally reported (expected factor structure) with the one obtained in this study (observed factor structure) was further analyzed by the CFA. The analysis was based on a higher number of parameters than is commonly reported. The CFA results showed that the ISS factor structure observed in our study significantly differs from the factor structure originally reported. The discrepancies are higher than those reported in the studies of Fritz et al. (Fritz et al., 2005; Fritz et al., 2002). Only the discriminative validity of the scale factors showed values acceptable by the Fornel-Larker criterion, although not quite convincingly.

Thus we find there are enough reasons and empirical evidence to support modification of the original ISS instrument. Furthermore, this research is not characterised by the limitations that made Fritz et al. (2005) unconfident about their own findings. The research in Serbian context was conducted on a sufficiently large and almost completely ethnically homogeneous sample, and the analyses of the factor structure were performed on all items of the scale.

In the first step, our proposal of modification is focused on the "Interaction Attentiveness" factor. The factor had a markedly low overall reliability and the same was found for the items comprising it. This was also reported in both studies used for comparison (Fritz et al., 2005; Fritz et al., 2002) and in the study of Graf and Harland (2005). Exploratory factor analysis showed, as well, that this factor was insufficiently well determined (see Table 1). A content examination of the factor's items suggests that, in comparison to the other four factors with a clear affective orientation, it is a cognitively oriented

factor. Its definition – "the participant's effort to understand what is going on in intercultural interaction" (Chen & Starosta, 2000b, p. 9) – is closely related to cognition. Thus, the scale was modified by excluding this factor, i.e. all items which the authors originally attributed to, as well as the items found to saturate it in this study. Furthermore, four items which significantly saturate more than one factor, both by exploratory and confirmatory analyses, were also excluded. Two items (13 and 22) were attributed to other factors according to the PAF and CFA results. A content analysis did not reveal misplacement. The modified scale is thus comprised of 15 items with four factors: "Interaction Enjoyment", "Interaction Engagement", "Respect for Cultural Differences" and "Interaction Confidence". We consider these factors to efficiently represent the affective dimension of intercultural competencies and we believe that the content of the majority of items comprising the scale clearly indicates this (items 1, 3, 8, 9, 10, 12, 15 and 24). The remaining items share content between emotional and cognitive dimensions (items 2, 5, 13 and 18) or emotional and behavioral dimensions (items 21, 22 and 23). According to Chen and Starosta (2000a), the shared content of items comprising the scale does not represent a deficiency in the operationalization of intercultural sensitivity. Intercultural sensitivity is based on intercultural consciousness which promotes, by itself, interculturally competent behaviour. The three dimensions of intercultural competence – cognitive, affective and behavioral – are necessarily linked.

Analysis of items representing "the best candidates" for exclusion from the scale, suggests two regularities. There are several items with low reliability (items 6, 11, 14 and 19), which were probably partly misunderstood by Serbian participants so they chose neutral responses (or used some other strategy in such a situation). An example is the item 19: *I am sensitive to my culturally-distinct counterpart's subtle meanings during our interaction*. This can be the case because of language complexity or because of hidden intentions, for instance.

The other group of items consists of those whose meanings are linked to the two or three latent dimensions, i.e. factors – in the way in which the factors are determined by the scale constructors. The point is in those items which, translated into Serbian language, express different components of Chen and Starosta's intercultural sensitivity model. It was manifested through the PAF and CFA results (items 4, 7, 16 and 20). For example, item 7 (*I don't like to be with people from different cultures*) is supposed to be linked only to "Respect for Cultural Differences", but it is also significantly linked to the factor expressing positive and negative reactions towards communicating with people from different cultures (i.e. "Interaction Enjoyment").

The data from PAF and CFA suggest that significant number of items for the participants in Serbia have a somewhat different meaning in relation to how these items are interpreted by American participants. The back translation method did not provide a necessary equivalence of the instrument in two different contexts. One of the difficulties of this method is its attachment to the literal meaning of the words, and consequently, its limitation in providing conceptual understanding of the items (Greenholtz, 2005; Kristjansson, Desrochers, & Zumbo, 2003).

Suggested procedures for instrument adaptations are more complex and usually involve bilingual subjects – through every phase (Greenholtz, 2005) or for the test-retest procedure in order to determine concurrent validity of the instrument (and its reliability) (Kristjansson et al., 2003).

It is also possible that certain lingua-cultural differences hinder a satisfactory adaptation of the ISS without a significant change to the content of this scale. The difference between the high-context and low-context communication is considered to be the most important. Hall was the first to define this dimension of culture which differentiates societies in terms of the degree to which the information in interpersonal interaction are verbally coded and explicitly expressed (Hall, 1991/1998). USA is characterized by low-context communication which demands verbal messages of high accuracy. On the other hand, Serbia belongs to high-context cultures in which the meaning of verbal messages relies, to a higher extent, to the context of communication. Since the test situation doesn't provide information about the context, there is a significant probability that the participants in Serbia will interpret the same statement differently.

We assessed the performance of the modified scale in our sample by CFA and found it performed significantly better than the original. Most fit indices satisfied expected criteria. This was especially evident in parameters of global fit. All those parameters of global fit generally referenced to in relevant literature were in accordance with the requirements. The value of chi square remained statistically significant, but it is known that non-normality of the data (which is the case in this study) can increase this statistic (Russell, 2002). The chi-square value also depends on the sample size. In models with large samples, trivial differences often cause the chi-square to be significant (Tabachnick & Fidell, 2007).

Advantage of the modified version of the ISS is tested with respect to several psychometric properties. Sensibility of the modified scale is not better nor worse than sensibility of the original scale. Respondents usually estimate themselves with scores higher than average, so the variability is increased on the left side of the distribution. The internal consistency of the modified scale has the same value as the alpha coefficient of the original one, although it contains nine items less. In addition, most of the factors in the structure of the modified scale have fewer items and higher coefficient of internal consistency. The construct validity of the modified ISS in relation to the CQS subscales proved to be good. And finally, it can be concluded that the modified ISS has better predictive power. It is more successful in the prediction of relevant criteria (given by the order of its significance): the CQS global score, the score on the motivational subscale of the CQS, the score on the cognitive subscale of the CQS and the score on the behavioral subscale of the CQS.

Based on the obtained results we believe that the modified ISS is more compatible with the conceptualisation of intercultural sensitivity by Chen and Starosta (2000a) than the original ISS.

Conclusion

We recommend the modified 15-item version of ISS for application in Serbia as more parsimonious than the original ISS, as well as a sufficiently reliable and valid (with respect to the structural, construct and predictive validity) measure of intercultural sensitivity. However, it is advisable to recheck the psychometric properties of the original instrument, as well as our proposed modification, on different, and perhaps more comprehensive samples in Serbia. We predict that the "Interaction Attentiveness" factor will continue to perform poorly and that the performance of the instrument will be enhanced by excluding items contributing to this factor, owing to the confounding effect of the cognitive dimension.

We also recommend a possibility of implementing a less direct translation of the items which had insufficient reliability or split factor loadings. Ideally, translation should provide a clearer meaning of the item, which more fully corresponds to the meaning of the latent dimension underlying the item. However, at the moment it is still uncertain whether it is possible to achieve this without causing significant changes of the items and the scale in general. This paper illustrates that adaptation of instruments for their application in a new lingua-cultural milieu can be a delicate process due to various factors.

References

Ang, S., Van Dyne, L., Koh, C., Ng, K. Y., Templer, K. J., Tay, C., & Chandrasekar, N. A. (2007). Cultural Intelligence: Its Measurement and Effects on Cultural Judgement and Decision Making, Cultural Adaptation and Task Performance. *Management and Organization Review, 3*, 335–371. doi: 10.1111/j.1740–8784.2007.00082.x

Bennett, J. M., & Bennett, M. J. (2004). Developing Intercultural Sensitivity: An Integrative Approach to Global and Domestic Diversity. In D. Landis, J. M. Bennet, & M. J. Bennet (Eds.), *Handbook of Intercultural Training* (3rd ed.) (pp. 147–165). Thousand Oaks, CA: Sage Publications, Inc.

Bhawuk, D. P. S., & Brislin, R. (1992). The Measurement of Intercultural Sensitivity Using the Concepts of Individualism and Collectivism. *International Journal of Intercultural Relations, 16*, 413–436.

Chen, G.-M., & Starosta, W. J. (2000a). Intercultural Sensitivity. In L. A. Samovar, & R. E. Porter, (Eds.), *Intercultural Communication: A Reader*. (pp. 406–414). Belmont, CA: Wadsworth Publishing Company.

Chen, G.-M., & Starosta, W. J. (2000b). The Development and Validation of the Intercultural Sensitivity Scale. *Human Communication, 3*, 1–15.

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, *18*, 39–50.

Fritz, W., Graf, A., Hentze, J., & Möllenberg, A. (2005). An Examination of Chen and Starosta's Model of Intercultural Sensitivity in Germany and United States. *Intercultural Communication Studies, 14,* 53–65.

Fritz, W., Möllenberg, A., & Chen, G.-M. (2002). Measuring Intercultural Sensitivity in Different Cultural Context. *Intercultural Communication Studies*, *11*, 165–176.

Graf, A., & Harland, L. K. (2005). Expatriate Selection: Evaluating the Discriminant, Convergent and Predictive Validity of Five Measures of Interpersonal and Intercultural Competence. *Journal of Leadership and Organizational Studies, 11,* 46–62.

Greenholtz, J. F. (2005). Does Intercultural Sensitivity Cross Cultures? Validity Issues in Porting Instruments across Languages and Cultures. *International Journal of Intercultural Relations, 29*, 73–89. doi: 10.1016/j.ijintrel.2005.04.010

Hall, E. T. (1991/1998). The Power of Hidden Differences. In M. J. Bennett (Ed.), *Basic concepts of intercultural communication: A Reader* (pp. 53–67). Boston, MA: Intercultural Press.

Hammer, M. R., Bennet, M. J., & Wiseman, R. (2003). Measuring Intercultural Sensitivity: The Intercultural Development Inventory. *International Journal of Intercultural Relations, 27*, 421–443. doi: 10.1016/S0147–1767(03)00032–4

Kristjansson, E. A., Desrochers, A., & Zumbo, B. (2003). Translating and Adapting Measurement Instruments for Cross-Linguistic and Cross-Cultural Research: A Guide for Practitioners. *The Canadian Journal of Nursing Research, 35,* 127–142.

Matsumoto, D., & Hwang, H. C. (2013). Assessing Cross-Cultural Competence: A Review of Available Tests. *Journal of Cross-Cultural Psychology, 44,* 849–873. doi: 10.1177/0022022113492891

Petrović, D. & Komnenić, D. (2012). Provera metrijskih karakteristika i faktorske strukture Skale kulturalne inteligencije. *Knjiga rezimea: Merenje i procena u psihologiji* (str. 182). Beograd: Društvo psihologa Srbije.

Petrović, D., & Zlatković, B. (2009). Intercultural Sensitivity of Future Primary School Teachers. In N. Popov, C. Wolhuter, B. Leutwyler, M. Mihova, J. Ogunleye, & Z. Bekiroğulları (Eds.), *Comparative Education, Teacher Training, Education Policy, and Social Inclusion* (pp. 121–128). Sofia: Bureau for Educational Services.

Russell, D. W. (2002). In Search of Underlying Dimensions: The Use (and Abuse) of Factor Analysis in Personality and Social Psychology Bulletin. *Personality and Social Psychology Bulletin, 28,* 1629–1646. doi: 10.1177/014616702237645

Starčević, J. (2013). *Kritička analiza konstrukta i provera konvergentno-diskriminativne valjanosti mera kulturalne inteligencije* (Unpublished doctoral dissertation). Filozofski fakultet, Beograd.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using Multivariate Statistics* (5th ed.). Boston, New York, San Francisco: Pearson Education. Inc.