

The development of the Humour Structure Appreciation Scale and its relation to Sensation Seeking Inventory and Need for Closure Scale

Filip Sulejmanov

Institute of Psychology, Ss Cyril and Methodius University, Skopje
filip.sulejmanov@gmail.com

Ognen Spasovski

Institute of Psychology, Ss Cyril and Methodius University, Skopje
ognen.spasovski@fzf.ukim.edu.mk

Tracey Platt

Institute of Psychology, University of Wolverhampton, United Kingdom
T.Platt@wlv.ac.uk

Abstract

Three related studies are presented with the aim of investigating the individual differences in humour appreciation, sensation seeking and need for closure. In order to accomplish this, the construction and initial validation of a new instrument will be presented. The Humour Structure Appreciation Scale (HSAS) distinguishes between incongruity-resolution (INCRES), and nonsense (NON) humour. The first study analyses the content validity of a broad item pool. The second study focused on the metric properties of the selected items. In the third study, the remaining jokes were given to a quota sample from Macedonia, along with measures of sensation seeking (Arnett Inventory of Sensation Seeking) and need for closure (Need for Closure Scale). Psychometric properties of the scales were investigated and showed to be satisfactory. Intercorrelations among the scales provide evidence for the construct validity of the HSAS.

Keywords: humour appreciation; sensation seeking; need for closure; HSAS.

1. Introduction

The appreciation of humour, as an individual difference construct, is well documented (Terry & Ertel 1974; Mindess et al. 1985; Ruch & Hehl 2007). Replicating the categories of humour derived from one country to another should confirm that the taxonomy is universal. Using

jokes that are comprehensive in a specific cultural context additionally should provide a valid instrument for the respective country.

From a series of factor-analytic studies (Ruch 1992) on a vast number of jokes and cartoons, three factors or dimensions of humour appreciation emerged, namely *incongruity-resolution* (INC-RES), *nonsense* (NON) and *sexual* (SEX) humour. These three dimensions of humour are operationalised in the 3WD (*Witz-dimensionen*) humour test (Ruch 1983). The first two factors relate to the structural properties of humour. INC-RES is defined as a two-stage process where the recipient first discovers an incongruity, which is then fully resolvable upon consideration of information available elsewhere in the joke or cartoon (Ruch 2008: 48). The conceptualised definition of NON acknowledges that the humour stimuli has an incongruous punch line, but “the punch line may 1) provide no resolution at all; 2) provide a partial resolution (leaving an essential part of the incongruity unresolved), 3) or create new incongruities” (McGhee et al. 1990: 124). The third factor, labelled SEX humour, refers to jokes and cartoons containing sexual themes. The response mode, based on factor-analytic studies, comprises two orthogonal and unidimensional components: *funniness* (f) and *aversiveness* (a).

Psychometric properties of the different forms of the 3WD test (50-item version, form K; and two parallel 35-item version, form A and B) are satisfactory, in terms of internal consistency, parallel test reliability and factor structure (for a review, see Ruch 1992). Although funniness of the structural dimensions of humour have a tendency to be positively correlated, factor-analytic studies showed that an orthogonal solution provided the best fit (Ruch 1992: 44). The extracted humour factors showed in many countries (Belgium, England, France, Germany, Israel, Italy and Turkey; Ruch & Hehl 2007).

Carretero-Dios et al. (2009, 2010) created a new instrument (also using jokes and cartoons), the Humour Appreciation Scale¹, comprising the three factors from the 3WD test, but expanding the content categories to include the content factors of *black* humour, *men disparagement* and *woman disparagement* humour. Carretero-Dios et al. (2006) postulate that “content should be isolated from structure (or cognitive processes) and the different structures and contents should be separated from each other” (as cited in Carretero-Dios et al. 2009: 782). The development of their humour test is aimed at accomplishing this dimensional proposal, but also to provide a methodological framework, or necessary stages for a construction of such a test. These stages, namely *content validity* and assessment of the metric properties of the items, will also be followed in our study.

Psychometric properties of the EAHU (content validity, metric properties of the items, exploratory factor structure and reliability of each dimension) are satisfactory and thoroughly presented in Carretero-Dios et al. (2009). It is interesting to note that in the first study (Carretero-Dios et al. 2009) a principal component analysis with Varimax-rotation is used, where in the subsequent study (Carretero-Dios et al. 2010) an oblique rotation is presented. Furthermore, Carretero-Dios et al. (2010) have employed Confirmatory Factor Analysis and provided additional evidence for the six factor solution.

Another measure was developed by Galloway and Chirico (2008) with a narrow focus on the structural dimensions of humour. The authors used only verbal jokes that were judged by the experimenters as good exemplars of NON and INC-RES humour. Separation between INC-RES and NON humour was shown with a principal component analysis using a Varimax rotation, since “an oblimin rotation indicated that none of the components were correlated with each other greater than +/- .30” (Diekhoff 1992, as cited in Galloway & Chirico 2008: 134). This instrument was developed to examine the association of humour appreciation with neuroticism.

How humour appreciation relates to personality traits has been a thoroughly examined question (for a review, see Ruch & Hehl 2007; Martin 2007) For the two structural

dimensions, variables like conservatism and agreeableness relate to having a preference for INC-RES, while sensation seeking and openness to experience relate to preferring NON.

A set of studies aimed at developing a new instrument that focuses on the two structural dimensions, the Humour Structure Appreciation Scale (HSAS), is presented. The first study entails the item construction and investigations of content validity. Here the aim was to select representative Macedonian jokes and to check “whether the domain of content for the construct is adequately represented by the items” (Haynes et al. 1995, as cited in Carretero-Dios et al. 2009: 775), focusing on verbal jokes. In the 3WD, all the NON items are cartoons. However, the NON in the EAHU is represented by both verbal jokes and cartoons. To our knowledge, only one study (Galloway & Chirico 2008) uses only verbal jokes for the two structural dimensions.

The rationale for the creation of the HSAS is to provide a valid instrument for investigating humour appreciation in Macedonia. Additionally, we were interested in whether the two structural factors would emerge if only verbal jokes that circulate in a specific cultural context are used. Jokes were selected from several sources, like web portals, newspapers and from students. The separation of the two structural factors using jokes that are comprehensible and relevant in our country would provide further evidence about the universality of structural properties of humour. Furthermore, nonsense humour is not usually encountered in jokes (Martin 2007: 205), so one of the reasons that we focus only on verbal jokes is to examine whether there would be a distinction between INC-RES and NON when such stimuli are used.

The second study investigates the psychometric properties of the constructed items with a pilot item analysis. Study 3 replicates the findings from the previous study while exploring the internal structure of the HSAS using principal component analysis (PCA). Here, we also investigate the construct validity of the HSAS by relating it to sensation seeking (Arnett 1994; Haynes, Miles & Clements 2000) and need for closure (Kruglanski & Webster 1996; Roets & Van Hiel 2011). The internal structure of these instruments is also explored, since this is the first time they have been used in our country. Nevertheless, the results will be focused on the HSAS.

Arnett’s (1994) conceptualisation of sensation seeking includes the dimensions of novelty and intensity of stimulation and experience. Carretero-Dios & Ruch (2010) showed that the novelty dimension was predictive of low appreciation of INC-RES humour and high appreciation of NON humour. However, these authors state that evaluating how universal the role of sensation seeking in humour appreciation is still necessary and that samples from non western countries should be included (Carretero-Dios & Ruch 2010: 431).

Need for closure is defined as “the desire to an answer on a given topic, any answer...compared to confusion and ambiguity (Kruglanski 1990: 337). To our knowledge, need for closure has not yet been related to humour structure appreciation. This construct is theoretically related to intolerance of ambiguity (Brenkelmann & Brenkelmann 1960; Kischkel 1984). Previous studies (Ruch & Hehl 1983; see Ruch 1992, for a review) showed that intolerance of ambiguity is related to higher funniness ratings of INC-RES humour and higher aversiveness ratings of NON humour. Ruch (1992) concludes that this variable is not a pure predictor of humour structure appreciation. However, need for closure is a much broader cognitive-motivational construct that could be relevant for explaining individual differences in humour structure appreciation. It is reasonable to assume that individuals high in need for closure should prefer INC-RES humour over NON.

2. Instrument Development Method

2.1. Study 1 Sample

The sample consisted of seven expert raters. Raters were all university professors within the field of psychology and with experience in the construction/adaptation of evaluation instruments.

2.2. Study 1 Instruments

A selection of 70 jokes were compiled into a test booklet. Each booklet also contained the instructions of the task (see Appendix A) and the conceptual definitions for the two structural dimensions of humour (see Appendix A).

2.3. Study 1 Procedure

A broad pool of around 200 jokes was selected by the first author from the Internet, newspapers and from students, who were given the conceptual definitions of INC-RES and NON humour. Afterwards, a group of three psychology students were asked to select jokes that best represent the conceptual definitions of both humour structure dimensions. In this way, an initial pool of 70 jokes (35 INC-RES and 35 NON) was proposed. The students were instructed to choose this number of jokes as similar number per dimension was used in Carretero-Dios et al. (2009) study. Also, they all had to agree on which dimension the joke corresponded to.

After an acceptance to cooperate, the experts were independently given the test booklet to complete, with different order of the jokes in each booklet. They were instructed to indicate, for each joke, how prototypical it was for either INC-RES or NON.

A Content Validity Index (CVI; Wynd et al. 2003) was used to analyse responses. A criterion for acceptance was that at least 5 out of the 7 judges nominated the joke as INC-RES or NON (see Carretero-Dios et al. 2009).

2.4. Results

The results indicated 19 jokes should be eliminated ($CVI < .70$). From the remaining jokes, 27 were categorised as INC-RES and 24 as NON (CVI of .71, 18 jokes, .86, 22 jokes, 1,11 jokes).

3. Pilot Item Analysis Method

3.1. Study 2 Sample

The overall sample consisted of 221 participants. The mean age was 33.03 years and ranged between 15 and 73 years, and $SD = 9.26$. The sample consisted of 45.7% men. Four versions of the instrument, with randomly allocated items in each, were given to the participants. The first was completed by $n = 55$ participants (29 men, $M = 34.02$; $SD = 6.81$, ranging 21 to 47 years). The second: $n = 55$ (33 men, $M = 36.40$; $SD = 11.16$, ranging 19 to 73 years). The third: $n = 55$ (16 men, $M = 28.98$; $SD = 7.89$, ranging 15 to 48 years), while the fourth was completed by $n = 56$ participants (23 men, $M = 32.71$; $SD = 9.09$, ranging between 21 to 61 years).

3.2. Study 2 Instruments

A total of 51 items (27 INC-RES and 24 NON) were used. The items in this study were rated only on the funniness component, from 0 = not funny at all, to 6 = very funny, since the Google forms platform doesn't allow two rating scales to be included. The aversiveness rating scale was not considered.

3.3. Study 2 Procedure

The study was conducted over the Internet on a Google forms platform that did not allow submission if items were missing. Four versions of the instrument (the 51 jokes randomly allocated for each version) plus 4 warm-up jokes, which were subsequently excluded from analysis, were given to participants. The inclusion of "warming up" items is consistent with the 3WD test (Ruch 1992).

3.4. Results

3.4.1. Analysis of variance

Two ANOVAs were conducted. In the first, the funniness score of INC-RES was considered as a dependent variable, while the fixed factors were the different groups of participants that completed the different versions of the instrument. In the second analysis, the same fixed factors were used, while the dependent variable was the funniness score of NON. Both analyses showed no statistically significant differences between the different groups of participants, allowing the four different subject groups of participants to be considered as one sample. All further statistical analyses are conducted using the total sample of 221 participants.

3.4.2. Descriptive statistics and item analysis

Out of the 51 initial items, 10 were selected (5 INC-RES, 5 NON). We used the criteria for eliminating items proposed by Carretero-Dios et al. (2009: 777): a) response range < 3; b) standard deviation < 1; c) corrected item-total correlation < .30; d) difference between the item scores of participants situated in the lower limit (percentile 27) and higher limit (percentile 73) of each dimension, positive or equal to 0. None of these analyses led to eliminating items. The only analysis that led to discarding items was the homogeneity criterion proposed by Jackson (as cited in Carretero-Dios et al. 2009: 777) that rules out elements (jokes) whose correlation with the total value of their dimension did not exceed their correlation with any of the other dimension by at least two tenths. The homogeneity analysis criterion led to eliminating 40 items. It should be noted that this analysis also led to discarding the greater number of items in Carretero-Dios et al. (2009) study.

Table 1 shows the descriptive statistics and discrimination indexes (difference scores between "high" and "low" groups in their funniness scores and CITC). Only the items that had satisfied the two-tenths criterion are given.

Table 1

Descriptive statistics and discrimination indexes of INC-RES and NON items

Items	M	SD	D	CITC	r
Incongruity resolution					
INC-RES1	3.10	1.86	-4.83	.69	.57**
INC-RES2	3.00	1.78	-4.13	.74	.58**
INC-RES3	2.66	1.87	-4.36	.78	.66**
INC-RES4	3.24	1.84	-4.90	.71	.58**
INC-RES5	3.18	1.73	-4.26	.74	.60**
Nonsense humour					
NON1*	1.89	1.87	-4.02	.34	.21*
NON2	2.33	1.98	-4.43	.57	.46**
NON3	2.58	1.98	-4.56	.66	.55**
NON4	1.63	1.84	-5.01	.52	.38**
NON5	2.47	2.01	-4.69	.55	.40**
NON6	3.17	2.01	-2.78	.70	.58**

Notes: $N = 221$, INC-RES = incongruity-resolution; NON = nonsense, D = difference between the mean scores between “high” and “low” groups (27th and 73rd percentile); CITC = corrected item-total correlation; r = correlation of the item with its own dimension.

*This item was eliminated from further analysis.

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

Table 1 shows that 11 items had at least a .20 higher item-total correlation with their own dimension in comparison with the other dimension. One NON item satisfying this criterion was eliminated, due to its rather small CITC value (.34)

4. Method – Study 3

Item analysis and PCA of a first proposal of the Humour Structure Appreciation Scale and its relation to sensation seeking and need for closure

4.1. Study 3 Sample

A quota sampling based on the region of living in Macedonia (West, East, Central and the capital city of Skopje), gender (similar number of men and women) and age groups (18 – 25, 26 – 40 and, 41 – 60 years) was utilised. The structure is shown in Table 2.

Additionally, information about participant’s level of education was included. In the sample, 2.3% of the participants had primary education, 52.5% had secondary education, 36.9% had a bachelor’s degree, and 8.3% held a MA degree.

Table 2

Structure of the sample used by region and gender

Gender	Region				
	West	East	Central	Skopje	Total
Men	23	30	18	41	112
Women	23	28	16	38	105
Total	46	58	34	79	217

Notes: The western region included participants from the cities of: Bitola, Ohrid, Struga and Kichevo; Eastern region was represented by the cities of: Kumanovo, Strumica, Vinica and Kratovo. The cities of Prilep and Veles were taken to represent the central region, while the capital city of Skopje was considered as a separate region.

4.2. Study 3 Instruments

Humour Structure Appreciation Scale (HSAS; Sulejmanov 2016) A first proposal of the scale was used with the selected items from the pilot item analysis - Study 2. The scale consisted of 10 items (5 INC-RES and 5 NON) with 4 “warm-up” items. The order of the items in the scale was randomly decided. Participants rate each item on two unipolar 7-point scales for “funniness” (from 0 = not at all funny, to 6 = very funny) and “aversiveness” (from 0 = not at all aversive, to 6 = very aversive).

Macedonian version of the Arnett Sensation Seeking Inventory (ASSI-M; Haynes et al. 2000). The back-translated from Macedonian to English scale consists of 11 items² (six from the intensity subscale, five from the novelty subscale). Participants indicate which response best applies to them for each item (A = describes me very well, to D = does not describe me at all). The inventory contains items such as, “I can see how it would be interesting to marry someone from a foreign country” (Novelty), “When I listen to music, I like it to be loud” (Intensity).

A principal component analysis with oblique rotation was performed for the AISS. The KMO value was 0.78, and Bartlett’s test showed statistical significance ($Chi-square = 438.77$, $df = 55$, $p < .01$). PA led to an estimate to retain two components. The two factor solution explained 42.0% (Eigenvalues were 3.27, 1.35, 1.11 and 0.99) of the variance. All of the items had loadings in the expected direction and above .30 (range between .42 to .81 for the intensity subscale; and .38 to .79 for the novelty subscale), with the exception of one intensity item, which had high secondary loading of .34. Cronbach alpha coefficients were .75, .65, and .70, for the Total Scale, Intensity, and Novelty, respectively.

Macedonian version of the Need for Closure Scale (NFC-M; Roets & Van Hiel 2011). Participants completed the 15-item version of the NFC (Webster & Kruglanski 1994; revised by Roets & Van Hiel 2007, 2011). The scale was back-translated from Macedonian to English. It is a one-dimensional measure of the need for closure construct. All items are positively keyed and use a six-point answer scale (from 1 = completely disagree, to 6 = completely agree). The scale contains items such as “I don’t like situations that are uncertain”, “I enjoy having a clear and structured mode of life”.

The psychometric properties of the scale were satisfactory. PA led to an estimate of one component to retain. All of the items had factor loadings above .30 (with an exception of one item which had a factor loading of .23). Cronbach alpha coefficient for the NFC was .81.

4.3. Results

4.3.1. Descriptive statistics and discrimination indexes

Descriptive statistics and corrected item-total correlations for funniness and aversiveness responses of the HSAS are given in Table 3.

Table 3

Descriptive statistics and item-total correlation for the HSAS

Items	Funniness			Aversiveness		
	M	SD	CITC	M	SD	CITC
HSAS						
Incongruity resolution						
INC-RES1	3.91	1.73	.39	0.98	1.47	.34
INC-RES2	4.25	1.81	.47	1.09	1.59	.47
INC-RES3	3.55	1.97	.50	1.31	1.71	.51
INC-RES4	3.82	1.95	.53	1.39	1.76	.52
INC-RES5	3.89	1.88	.58	1.12	1.64	.49
Nonsense humor						
NON1	1.86	1.82	.46	1.69	2.00	.66
NON2	2.49	1.84	.42	1.41	1.73	.47
NON3	2.02	2.11	.45	1.95	2.22	.51
NON4	2.60	2.16	.55	1.67	2.05	.63
NON5	2.60	2.16	.55	1.47	1.88	.51

Notes: $N = 217$; INC-RES = incongruity-resolution; NON = nonsense; CITC = corrected item-total correlation for each domain.

Table 3 shows that the corrected item-total correlations for INC-RES and NON items are all above .30.

4.3.2. Internal structure of the HSAS

First, a principal component analysis with oblique rotation was performed for the HSAS by funniness scores. The KMO value was 0.79, and Bartlett's test showed statistical significance ($Chi-square = 499.34$, $df = 45$, $p < .01$). Parallel analysis (PA) was applied in order to determine the number of components to retain (e.g., O'Connor 2000; Zwick & Velicer 1986). The PA was carried out based on Patil, Singh, Mishra & Donovan's (2007) engine and it led to an estimate of two components to retain. The two-factor solution explained 48.7% (Eigenvalues were 3.40, 1.47, 1.12 and 0.79) of the variance. Factor loadings of the 10 items are presented in Table 4.

Table 4

Loadings of the HSAS items after oblique rotation

Items	FI	FII	h ²
Incongruity resolution			
INC-RES1	.53	.14	.30
INC-RES2	.68	-.04	.46
INC-RES3	.62	.14	.40
INC-RES4	.78	-.09	.62
INC-RES5	.78	-.02	.61
Nonsense			
NON1	.12	.59	.36
NON2	.14	.55	.32
NON3	-.21	.77	.64
NON4	.03	.72	.52
NON5	.07	.73	.54

Notes: $N = 217$; INC-RES = incongruity-resolution; NON = nonsense; h^2 =communality.

Table 4 shows that INC-RES items have factor loadings above .40 on the first factor (range between .53 to .78) and no secondary loadings higher than .30. NON items also have high factor loadings above .40 (range between .55 to .73) on the second factor and none of the items have high secondary loadings on the first factor.

4.3.3. Scale reliabilities and intercorrelations

Intercorrelations among the measures used and Cronbach's alphas were computed. The results are given in Table 5.

Table 5
Intercorrelation between age, education and HSAS, NFC and AISS, and Cronbach's alphas

	INC-RESf	NO Nf	SPI	INC-RESa	NO Na	SsTotal	I	N	NFC	Age	Edu
INC-RESf	<i>.73</i>	<i>.41*</i>	<i>.50*</i>	<i>-.38**</i>	<i>-.12</i>	<i>-.11</i>	<i>-.06</i>	<i>-.12</i>	<i>.07</i>	<i>.08</i>	<i>.16*</i>
NONf		<i>.72</i>	<i>-.59**</i>	<i>-.09</i>	<i>-.30**</i>	<i>.11</i>	<i>.09</i>	<i>.10</i>	<i>-.20**</i>	<i>-.18**</i>	<i>.10</i>
SPI			<i>--</i>	<i>-.25**</i>	<i>.18*</i>	<i>-.20**</i>	<i>-.14*</i>	<i>-.21**</i>	<i>.26*</i>	<i>.25*</i>	<i>.05</i>
INC-RESa				<i>.71</i>	<i>.55*</i>	<i>-.02</i>	<i>.00</i>	<i>-.05</i>	<i>-.02</i>	<i>-.03</i>	<i>.00</i>
NONa					<i>.78</i>	<i>-.14*</i>	<i>-.06</i>	<i>-.18**</i>	<i>.10</i>	<i>.11</i>	<i>-.05</i>
SsTotal						<i>.75</i>	<i>.87*</i>	<i>.83*</i>	<i>-.32**</i>	<i>.41**</i>	<i>-.13</i>
I							<i>.65</i>	<i>.43*</i>	<i>-.31**</i>	<i>.32**</i>	<i>.17*</i>
N								<i>.70</i>	<i>.24**</i>	<i>.37**</i>	<i>-.05</i>
NFC									<i>.81</i>	<i>.32*</i>	<i>-.07</i>

Notes: N = 217

INC-RES = incongruity-resolution; NON = nonsense; SPI = structure preference index (INC-RESf – NONf); Ss = sensation seeking; I = intensity subscale; N = novelty subscale; NFC = need for closure; f = funniness; a = aversiveness; Edu = Education. Cronbach's alphas are italicised in the diagonal.

*p < .05; **p < .01

Table 5 shows that Cronbach's alphas ranged from .65 (intensity subscale) to .81 (need for closure scale). Although acceptable, the reliabilities of INC-RESf, NONf, and INC-RESa are low.

Socio-demographic variables were related to humour structure appreciation, sensation seeking, and need for closure. The level of education went along with higher scores in INC-RESf. Younger participants found NON funnier than did older participants. The same relation was found when the *structure preference index* (SPI) was taken into account. Age related negatively with sensation seeking total score, intensity and novelty, while the level of education was negatively related with intensity subscale scores. There was a positive correlation between age and need for closure.

Table 5 shows that the relation between sensation seeking and humour structure appreciation was in the expected direction and confirms previous findings. The SPI was negatively associated with sensation seeking total score, and both subscales. Regarding the aversiveness ratings, only higher novelty scores went along with lower scores in aversiveness of NON.

Furthermore, the expected pattern of relationship between need for closure and humour structure appreciation was confirmed, at least for the funniness ratings. Need for closure correlated negatively with NONf, and there was a positive association with the SPI. Furthermore, need for closure was negatively related with sensation seeking total score, and its components.

Finally, funniness of INC-RES and NON was positively related. INC-RESf and INC-RESa correlated negatively. This was also true for the nonsense humour components and is comparable with the 3WD and EAHU.

4.4. Discussion

The discussion follows the order of the presented studies. First, in the item-construction stage a representative sample of verbal jokes from Macedonia was selected. Of course, this does not mean that the origin of the jokes is Macedonian, but that the jokes fit a Macedonian context. Ensuring that the sample is heterogeneous—in the sense that the jokes are applicable for assessing a broad age range—several sources for collecting the jokes were utilised as it is a reasonable assumption that web portals have younger audience in comparison to newspapers. The web portals that were considered are among the top 10 visited in Macedonia and we considered almost all of the daily newspapers. Additionally, jokes were collected from students, who were also asked to tell jokes they heard from their parents and older siblings and peers. In this way, jokes that circulate orally were also included. From 200, 70 jokes were chosen for the content validity study.

Addressing the decision to focus only on verbal jokes; NON can be usually encountered in cartoons, literature and films instead of canned jokes (Martin 2007: 205). In Ruch's 3WD test all of the NON stimuli are cartoons. Our aim was to investigate whether the NON factor emerges if only verbal jokes are used. Although the cognitive processes that are triggered when humorous material is perceived are likely universal (Ruch & Köhler 2007: 208), it is reasonable to assume that in some cultures (perhaps the ones with strong uncertainty avoidance index) (Hofstede et al. 2010) the NON factor may not emerge if only verbal jokes are considered. According to Hempelmann & Samson (2008), there is an overlap between verbal and visual humour in regard to the cognitive processes, but there are some crucial differences in cognitive and aesthetic characteristics, which can change the affective response to the cartoon.

The content validity is often a stage that is not considered in creating a new psychological instrument, in spite of the fact that is an important step (Gordon 2004; Smith et

al. 2003; Sireci 1998, as cited in Carretero-Dios et al. 2009: 781). This analysis allowed the elimination of 19 jokes. Although the content validity study is a necessary first step, it is definitely insufficient for the creation of a valid psychological instrument. The next studies were aimed at obtaining empirical evidence for the distinction of INC-RES and NON in Macedonia.

As question order effects (Bradburn & Mason 1964; Feldman & Lynch 1988) could be relevant to humour appreciation the pilot item analysis checked whether the order of the items relates to INC-RES and NON score. No statistical differences were found so the further analysis was done using the total sample of 221 subjects.

The 10 items (5 INC-RES and 5 NON) were selected with satisfactory metric properties, the only relevant criterion being the one proposed by Jackson (1970), and were analysed using only the funniness scores. One limitation of this study is that we have not provided information about the variability in the aversiveness response mode. The reason for this was the Internet software used for collecting the data was Google forms. The Carretero-Dios et al. (2009) study consisted mainly of university students. Yet, a diverse sample should be used. Samples collected via Internet elicit diverse as well as reliable and valid participation (Gosling et al. 2004).

In the third study, we used a quota sample based on region of living in Macedonia, age and sex, although the quotas of the relevant characteristics are only approximate to the ones in the population. Additionally, the sample size was relatively small. We will first discuss the metric properties and internal structure of the HSAS and AISS, and afterwards, we will focus on the relation between humour structure appreciation, sensation seeking, and need for closure. It should be stressed that a main limitation of this study is that we used the Macedonian version of the AISS and NFC scales for the first time. It would have been desirable to establish the psychometric properties of these scales separately but available resources did not allow this.

The metric properties of the items from the HSAS obtained in the previous study were confirmed. The two factor solution was shown using a parallel analysis which is one of the most strongly recommended techniques for determining the number of components to retain (Courtney 2013: 4). All of the factor loadings were satisfactory and the factors could be easily interpreted as INC-RES and NON. An oblique rotation was appropriate since these two dimensions were positively correlated.

The AISS was first translated from English to Macedonian by the first author and back-translated to English by the second author. We should point that we have not encountered any problematical phrases. All of the factor loadings were satisfactory and in the correct direction, except for one intensity item which had high secondary loading on the first factor. However, it should be stressed that the intensity subscale had a rather low internal consistency. The AISS (Arnett 1994) had problematical psychometric properties in the study done by Carretero-Dios & Ruch (2010), where the alphas were low. Although we have used the modified shortened version proposed by Haynes et al. (2000), which is considered as an appropriate measure of sensation seeking, there was a low alpha for the intensity subscale in our sample. Concerning this, interpretation of the results related to sensation seeking should be interpreted with caution.

Regarding the correlations between the scales, we should first note that the structural dimensions of humour correlated positively ($r = .41$), which is almost identical to the finding reported by Carretero-Dios et al. (2010), where the correlation was .40, and comparable to the results Ruch reported (between $r = .28$ and $r = .38$). Contrary to the findings reported by Carretero-Dios & Ruch (2010), there was no significant relation between sensation seeking and funniness scores of INC-RES and NON. However, Carretero-Dios & Ruch (2010: 443), postulate that sensation seeking should be understood as a determinant for the relative

preference of one humour type over the other, and this is best assessed by the structure preference index. We have found a significant negative relation between the SPI and sensation seeking. The relation was also significant with intensity and novelty subscales. Individuals who don't like novel and intense stimulations prefer jokes that have fully resolvable incongruities in comparison with nonsense humour jokes which are more abstract.

As for the relation between sensation seeking and aversiveness scores of both structural dimensions, the results are comparable to earlier findings (Carretero-Dios & Ruch 2010). The results replicate their findings of a negative correlation between NONa and AISS total score and novelty subscale. The aversiveness response mode is related to the negative evaluations of humour (boring, aggressive, offensive etc.). This finding is further evidence that sensation seekers (novelty seekers) prefer NON over INC-RES.

The hypothesised relation between need for closure and humour structure appreciation was confirmed for the funniness response mode. Individuals with a high need to achieve closure, or in other words, an answer to a given topic compared to confusion and ambiguity preferred INC-RES humour over NON. This was confirmed with the negative relation with NONf, and positive relation with the SPI. It could be argued that INC-RES humour provides an answer (full resolution), or a sense that there is a point that has been understood. On the other hand, as nonsense humour is more bizarre, it creates confusion and ambiguity in individuals with a strong need to attain closure. Future studies could use the full revised version of the NFC scale (Roets & Van Hiel 2007), which enables assessment of the individual NFC facets. This could further illuminate the relationship between humour structure appreciation and need for closure. For example, is the NFC construct a predictor for preference of residual incongruities in humour, over and above intolerance of ambiguity?

Humour structure appreciation was also related to some of the socio-demographic variables. A counter intuitive finding was the positive relation between level of education and funniness of INC-RES humour. According to Ruch (1999: 92), enjoyment of NON comes from play with our ability and tendency to make sense and it can be considered as a "higher" form of humour. This would imply that the positive relation should be found between level of education and NONf, rather than INC-RESf. Future studies should examine the relation of education and humour structure preference in a more detail. The type of education should also be considered. Education can open our minds to play with ideas, but also can narrow this tendency and provide firm answers.

Finally, an expected relation was found between age and preference of NON. Carretero-Dios et al. (2010: 323) states this is due to nonsense humour being related with sensation seeking and this predictor decreases with age. Our results showed that older participants decline in sensation seeking in general, but this trend was also true for novelty and intensity.

The next step should be replicating these results on a greater sample and applying confirmatory factor analysis. The convergent validity of the HSAS should be confirmed using the 3WD and EAHU. A future study could involve a joint factor analysis of the humour stimuli from the structural dimensions from the three scales. Also, other personality variables should be taken into account. Additionally, we should explore which content categories will be relevant to our culture. A possible candidate for a content category in Macedonia (and Balkan countries) is *political* humour. Mindess et al (1985) distinguish ten categories of humour: *nonsense, philosophical, sexual, scatological, social satire, hostile, demeaning to men, demeaning to women, ethnic and sick*. However, we advise that if the aim is to determine the relevant content categories of humour in a specific culture, an exploratory approach is more adequate than deriving the content categories on a basis of theoretical frameworks.

5. Conclusions

The HSAS showed satisfactory psychometric properties and initial validation is provided. The results are consistent with previous research regarding the separation of incongruity resolution and nonsense humour. Cross-cultural support is also provided for the relation of sensation seeking and humour structure appreciation.

Acknowledgements

The authors thank Dr. Sonja Heintz for helpful comments.

Notes

The original, Spanish, name of the scale is *Escala de Apreciación del Humor* (EAHU).

² See pp. 830 (model A2b) in Haynes et al. (2000) for the instrument. We have not included items 2 and 7, since their factor loadings are below .30.

Appendix A

INSTRUCTIONS FOR JUDGES/EXPERTS IN THE CONTENT VALIDITY TASK

On the following pages there are 70 jokes. In humor appreciation there is a difference between the cognitive processes that are triggered when perceiving jokes, or two types of humor. These types of humor are labeled: incongruity-resolution (INC-RES) and nonsense (NON) humor.

In INC-RES humor, a two-stage process can be observed: the perception of an incongruity and its resolution. In this type of humor, an incongruity is discovered and then resolved using the information available elsewhere in the joke. In other words, the common element in this type of humor is that the recipient first discovers an incongruity which is then fully resolvable upon consideration of information available elsewhere in the joke. There is a certain projective element in these jokes as essential things are not spelled out and have to be supplemented by the recipient; often resolving the incongruity requires attributing motives and traits (e.g, stingy, mean, stupid, absent-minded) to the characters depicted in the jokes. Although individuals might differ with respect to how they perceive and/or resolve the incongruity, they have the sense of having “gotten the point” or understood the joke once resolution information has been identified.

Nonsense humor also has a surprising or incongruous punch line, however, the punch line may (1) provide no resolution at all, (2) provide a partial resolution (leaving an essential part of the incongruity unresolved), or (3) actually create new absurdities or incongruities”. In nonsense humor the resolution information gives the appearance of making sense out of incongruities without actually doing so. The recipient’s ability to make sense or to solve problems is exploited; after detecting the incongruity he is misled to resolve it, only to later discover that what made sense for a moment is not really making sense.

Please in the box below each joke indicate in which humor type the joke belongs. If you think that the joke should be classified as incongruity-resolution, write **INC-RES** in the box bellow the joke. If you think that the joke should be classified as nonsense, write **NON** in the box bellow the joke.

References

- Arnett, J. (1994). 'Sensation seeking: a new scale and a new conceptualisation'. *Personality and Individual Differences*, 16(2), pp. 289-296.
- Bradburn, N. M., & Mason, W. M. (1964). 'The effect of question order on responses'. *Journal of Marketing Research*, 1(4), pp. 57-61.
- Brengelmann, J. C. & Brengelmann, L. (1960). 'German validation of questionnaires assessing dogmatic and intolerant attitudes'. *Zeitschrift für experimentelle und angewandte Psychologie*, 7, pp. 451-471.
- Carretero-Dios, H., Perez, C. & Buela-Casal, G. (2006). 'Dimensiones de la apreciación del humor [Dimensions of humor appreciation]'. *Psicothema*, 18, pp. 465-470.
- Carretero-Dios, H., Perez, C., & Buela-Casal, G. (2009). 'Content validity and metric properties of a pool of items developed to assess humour appreciation'. *The Spanish Journal of Psychology*, 12(2), pp. 773-787.
- Carretero-Dios, H., Perez, C., & Buela-Casal, G. (2010). 'Assessing the appreciation of the content and structure of humour: Construction of a new scale'. *Humour: International Journal of Humour Research*, 23, pp. 307-325.
- Carretero-Dios, H. & Ruch, W. (2010). 'Humour appreciation and sensation seeking: Invariance of findings across culture and assessment instrument?' *Humour: International Journal of Humour Research*, 23, pp. 427-445.
- Courtney, M. G. R. (2013). 'Determining the number of factors to retain in EFA: Using the SPSS R-Menu v2.0 to make more judicious estimations'. *Practical Assessment, Research & Evaluation*, 18(8), pp. 1-14.
- Diekhoff, G. (1992). *Statistics for the Social and Behavioural Sciences: Univariate, Bivariate, Multivariate*. Dubuque, IA: Brown.
- Feldman, J. M., & Lynch, J. G. J. (1988). 'Self-generated validity and other effects of measurement on belief, attitude, intention, and behaviour'. *Journal of Applied Psychology*, 73(3), pp. 421-435.
- Galloway, G., & Chirico, D. (2008). 'Personality and humour appreciation: evidence of an association between trait neuroticism and preferences for structural features of humour'. *Humour: International Journal of Humour Research*, 21, pp. 129-142.
- Gordon, J. (2004). 'Developing and improving assessment instruments'. *Assessment in Education: Principles, Policy and Practice*, 11, pp. 243-245.
- Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). 'Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires'. *American Psychologist*, 59, pp. 93-104.
- Haynes, C. A., Miles, N. V., Clements, K. (2000). 'A confirmatory factor analysis of two models of sensation seeking'. *Personality and Individual Differences*, 29, pp. 823-839.
- Haynes, S.N., Richard, D.C.S., & Kubany, E.S. (1995). 'Content validity in psychological assessment: A functional approach to concepts and methods'. *Psychological Assessment*, 7, pp. 238-247.
- Hempelmann, C. F., & Samson, A. C. (2008). 'Cartoons: Drawn jokes?' In V. Raskin (Ed.), *A Primer of Humour Research*, Berlin: Mouton de Gruyter. pp. 609-40
- Hofstede, G., Hofstede, G. J., & Minkov M. (2010). *Cultures and Organisations: Software of the Mind: Intercultural Cooperation and its Importance for Survival*, 3rd ed., New York: McGraw Hill.
- Jackson, D.N. (1970). 'A sequential system for personality scale development', in C.D. Spielberger (ed.), *Current Topics in Clinical and Community Psychology* (vol. 2), New York: Academic Press. pp. 61-96

- Kischkel, K. H. (1984). 'A scale to assess tolerance of ambiguity'. *Diagnostica*, 30, pp. 144-154.
- Kruglanski, A.W. (1990). 'Motivation for judging and knowing: Implications for casual attribution'. In E.T. Higgins & R.M. Sorrentino (Eds.), *The Handbook of Motivation and Cognition: Foundation of Social Behaviour* (vol. 2), New York: Guilford Press. pp. 333-368
- Kruglanski, A.W., & Webster, D.M. (1996). 'Motivated closing of the mind: "Seizing" and "Freezing"'. *Psychological Review*, 103, pp. 263-283.
- Martin, R. A. (2007). *The Psychology of Humour: An Integrative Approach*. Burlington, MA: Elsevier Academic Press.
- McGhee, P.E., Ruch, W., & Hehl, F.J. (1990). 'A personality-based model of humour development during adulthood'. *Humour: International Journal of Humour Research*, 3, pp. 119-146.
- Mindess, H., Miller, C., Turek, J., Bender, A., & Corbin, S. (1985). *The Antioch Humour Test: Making Sense of Humour*. New York: Avon Books.
- O'Connor, B. P. (2000). 'SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test'. *Behaviour Research Methods, Instruments, & Computers*, 32, pp. 396-402.
- Patil, V. H., Singh, S. N., Mishra, S. & Donovan, T.D. (2007), Parallel Analysis Engine to Aid Determining Number of Factors to Retain [Computer software]. Available from <http://smishra.faculty.ku.edu/parallelengine.htm> (accessed August 8, 2016)
- Roets, A. & Van Hiel, A. (2007). 'Separating ability from need: Clarifying the dimensional structure of the need for closure scale'. *Personality and Social Psychology Bulletin*, 33(2), pp. 266-280.
- Ruch, W. (1983). 'Humour-Test 3WD (Forms A, B, and K)'. Unpublished manuscript, University of Dusseldorf, Germany.
- Ruch, W. (1992). 'Assessment of appreciation of humour: Studies with the 3WD humour test', in C. D. Spielberger & J. N. Butcher (Eds.), *Advances in Personality Assessment* (Vol. 9), Hillsdale, NJ: Erlbaum. pp. 27-75
- Ruch, W. (1999). 'The sense of nonsense lies in the nonsense of sense. Comment on Paolillo's (1998) Gary Larson's Far Side: Nonsense? Nonsense! – Debate'. *Humour: International Journal of Humour Research*, 12(1), pp. 71-93.
- Ruch, W. (2008). 'The psychology of humour', in Raskin V. (Ed.), *A Primer of Humour Research*, Berlin: Mouton de Gruyter. pp. 17-100
- Ruch, W., & Hehl, F.-J. (1983). 'Intolerance of ambiguity as a factor in the appreciation of humour'. *Personality and Individual Differences*, 4, pp. 443-449.
- Ruch, W., & Hehl, F.-J. (2007). 'A two-mode model of humour appreciation: Its relation to aesthetic appreciation and simplicity-complexity of personality', in Ruch, W. (Ed.), *The Sense of Humour: Exploration of a Personality Characteristic*, Berlin: Walter de Gruyter. pp. 109-142
- Ruch, W. & Köhler, G. (2007). 'A temperament approach to humour'. In W. Ruch (Ed.), *The Sense of Humour: Explorations of a Personality Characteristic*. Berlin: Mouton de Gruyter. pp. 203-231
- Sireci, S. G. (1998). 'The construct of content validity'. *Social Indicators Research*, 45(1-3), pp. 83-117.
- Smith, G. T., Fischer, S., & Fister, S. M. (2003). 'Incremental validity principles in test construction'. *Psychological Assessment*, 15(4), 467.
- Sulejmanov, F. (2016). 'Humour Structure Appreciation Scale'. Unpublished Instrument. University of Ss Cyril and Methodius University, Skopje, Macedonia.

- Terry, R. L., & Ertel, S. L. (1974). 'Exploration of individual differences in preferences for humour'. *Psychological Reports*, 34(3), pp. 1031-1037.
- Webster, D. M. & Kruglanski, A. W. (1994). 'Individual differences in need for cognitive closure'. *Journal of Personality and Social Psychology*, 67(6), 1049.
- Wynd, C.A., Schmidt, B., & Atkins-Schaefer, M. (2003). 'Two quantitative approaches for estimating instrument content validity'. *Western Journal for Nursing Research*, 25, pp. 508-518.
- Zwick, W. R., & Velicer, W. F. (1986). 'Comparison of five rules for determining the number of components to retain'. *Psychological Bulletin*, 99, pp. 432-442.