ASK, 2008, nr 17, strony 7–37

THE QUALITY OF THE MEASUREMENT OF INTEREST IN THE POLITICAL ISSUES PRESENTED IN THE MEDIA IN THE ESS

Lluís Coromina
Willem E. Saris
Daniel Oberski
Universitat Ramon Llull.
Barcelona (Spain)

It is simply impossible to know all about the world by one's own experience. Therefore, the media have an important role to play in informing people about events that occur and the interpretation of these events. However, people do not have to pay attention to this information. There is a significant difference between people's interest in a topic and following it in the media. Given the importance of the media for opinion building in Europe, the European Social Survey (ESS) included questions about the media in its questionnaire. In this study we concentrate on the measurement of "Interest in Political Issues presented in the Media". In the context of political orientation, this variable plays an important role. We evaluate the quality of the measurement instrument and the comparability of the measures across the European countries involved in the ESS. External validity of the instrument will be studied by examining whether relationships suggested in the literature for this measure are indeed found.
Key words: comparability, scalar invariance, metric invariance, composite scores.

Survey Research Centre, ESADE Business School, Universitat Ramon Llull. Av. Pedralbes, 60-62. 08006 Barcelona (Spain)

The basis for the IPIM measure

The aim of this paper is to study the quality of measurement for the "Interest in the Political Issues presented in the Media" -denoted by IPIM- and its comparability across countries participating in the European Social Survey (ESS) in three different rounds. The IPIM measure evaluated in this study is developed on the basis of the media use questions specified for the Core questionnaire of the ESS (Newton 2000). We begin with the introduction of this module. For a detailed discussion of the development of the media use module we refer to the ESS[1] website. The final proposal that has been used in the first three rounds of the ESS is presented in Figure 1.

Figure 1. Final questions for the media use module

A1 On an average weekday, how much time in total do you generally spend watching television?

No time at all	00
Less than ½ hour	01
½ hour to 1 hour	02
More than 1 hour, up to1½ hours	03
More than 1½ hours, up to 2 hours	04
More than 2 hours, up to 2½ hours	05
More than 2½ hours, up to 3 hours	06
More than 3 hours	07
(Don't know)	88

(filter if watch no television)

A2 And again on an average weekday how much of your time watching television is generally spent watching news or programmes about **politics and current affairs**?

Idem Scale

A3 On an average weekday, how much time in total do you generally spend listening to the radio?

Idem Scale

(filter if do not listen to radio)

A4 And again on an average weekday, how much of your time listening to the radio is generally spent listening to news or programs about **politics and current affairs**?

Idem Scale

A5 On an average weekday, how much time in total do you generally spend reading the newspapers?

Idem Scale

(filter if do not read newspapers)

A6 And how much of this time is generally spent reading about **politics and current affairs**?

Idem Scale

[1] http://www.europeansocialsurvey.org/.

Operationalization of Interest in Political Issues presented in the Media

The construct *Interest in Political Issues presented in the Media* (IPIM) is a subjective variable which should indicate how much importance a person gives to political issues presented in the media. One can argue that the importance which a person attributes to these issues is expressed in the amount of time he/she spends collecting this information from the media. This information has been asked in the questions A2, A4 and A6 (see Figure 1). A problem with this approach is that people may have available different amounts of time to spend on the media; for example a person who has a full time job can spend less time on the media than a person who is unemployed. So, the information from Figure 1 cannot be used directly to construct a measure for IPIM. A more valid measure can be obtained using a ratio of the amount of time spent on political issues in a medium over the total amount of time spent on this medium. Such a proportion can be specified for each type of medium. Three different specific measures for IPIM can thus be created. We chose to use the following measures:

$$IPIM_{TV} = \frac{A_2}{A_1}; \quad IPIM_{Radio} = \frac{A_4}{A_3}; \quad IPIM_{Newspaper} = \frac{A_6}{A_5}.$$

Here the A_i represents a transformation of the measures presented in Figure 1 to numerical scores, minutes in this case, which will be elaborated further below.

We assume that the construct or concept by postulation (Saris & Gallhofer 2007) IPIM can be operationalized by these three reflective indicators ($IPIM_{TV}$; $IPIM_{Radio}$; $IPIM_{Newspaper}$) which range from 0 to 1, where 0 means that respondents are not using the medium at all for political issues, and 1 means that they use the medium exclusively for political purposes. The model for this construct is presented in Figure 2 below.

The main purpose of this study is to create and compare an IPIM index which can be used in comparative research across countries and time. This index will be the sum of the observed variables ($IPIM_{TV}$, $IPIM_{Radio}$ and $IPIM_{Newspaper}$). The construction is presented in Figure 3.

Figure 2. The measurement model for IPIM

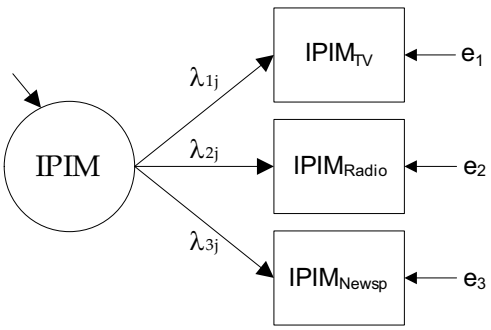
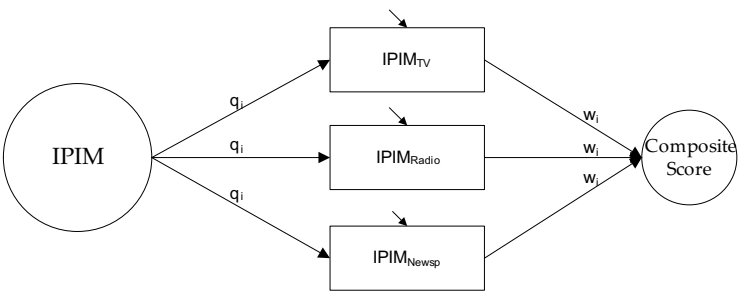


Figure 3. The construction of an index for IPIM as a composite score



In Figure 3 the q_i is the quality for each reflective indicator and w_i is the weight used for computing the composite score[2].

Because the IPIM scores for the different media are defined as ratios, their computation requires that the variables involved are numeric. This is not the case, however, as shown in Figure 1. In order to be able to compute numerical scores for the measure based on the "8 numerical categories" we must transform the original scale into a numerical one.

After this transformation we will compare the results of confirmatory factor analysis (CFA) models across different countries in the ESS testing for measurement invariance. Next we will estimate the quality of the IPIM index expressed in the correlation between the index and the latent variable. This will be carried out on the first and second ESS rounds.

Finally, substantive relationships will be studied on the first ESS round for 21

[2] The missing values for each IPIM were imputed with random scores with the same mean and standard deviation as the non-missing ones. This unconditional mean imputation with a random element is similar in idea to cold deck imputation (Little & Rubin 1987).

countries with some outcome variables which are related to IPIM and political interest according to our hypotheses derived from the literature.

Transformation of the categories

Each category (except the 8th category) is originally labelled with an upper and lower bound. The middle of the values is used for numerical transformation for all categories. For the last category, there is no upper bound available and it is unclear which value should be chosen. Fortunately, in the first ESS round we can make use of the additional questionnaire for computing the numerical values of this 8th category. In that round, open-ended questions about media use asked directly for a numeric estimate of the amount of time spent on each medium[3]. Using these variables a value is chosen for the last category in each country. Details of the whole procedure can be found at the website of the Survey Research Centre of ESADE[4]. The results were, for instance, that the numerical values obtained in Austria for the 8th category are 234 minutes for television, 278 for radio and 180 for newspaper. For Spain they are 261, 297 and 257 or for Belgium they are 269, 300 and 300, respectively.

Table 1 presents the transformation to minutes. The newly created numeric values of these variables can be found on the website of the Survey Research Centre[5] of ESADE.

Table 1. Original and transformed media use measures

Original Scale	Category	Transformation
00	No time at all	0 minutes
01	Less than ½ hour	15 minutes = ¼ hour
02	½ hour to 1 hour	45 minutes = ¾ hour
03	More than 1 hour, up to1½ hours	75 minutes = 1¼ hours
04	More than 1½ hours, up to 2 hours	105 minutes = 1¾ hours
05	More than 2 hours, up to 2½ hours	135 minutes = 2¼ hours
06	More than 2½ hours, up to 3 hours	165 minutes = 2¾ hours
07	More than 3 hours	Estimated mean for each medium

[3] The data can be obtained from the website http://www.europeansocialsurvey.org or http://ess.nsd.uib.no.

[4] http://www.esade.edu/research/src/investigacion/surveydesign.

[5] "Scores" in http://www.esade.edu/research/src/investigacion/surveydesign.

The values obtained in each country are used also in the 2nd ESS round. However, the participating countries in this ESS round were not exactly the same as the ones in the 1st round. For instance, Estonia participated in the 2nd but not in the first round; consequently additional information about these new countries was missing. However, we had to transform the categorical variables in numerical variables also for these countries. To solve this problem we specified for the 8th category for each new country the amount of minutes of a country from the 1st round which has the most similar distribution for each medium.

Comparability across countries

In order to be able to compare relationships and means across countries one has to show first that the measures in the different countries are comparable. In the ESS much time and resources have been spent to make the questions as comparable as possible across countries (Harkness et al. 2007) and to make the samples, as well, as comparable as possible (Lynn et al. 2007). Nevertheless, it has been shown by Saris & Gallhofer (2007), Saris & Oberski (forthcoming) and Oberski & Saris (2007) that this does not necessarily mean that the measures are equally good. Therefore, it is necessary to test what is called the "invariance of measurement instruments" across countries.

Invariance of the measurement instruments

The commonly[6] used criteria for measurement invariance are configural, metric and scalar invariance (Horn et al. 1983; Meredith 1993; Steenkamp & Baumgartner 1998). In order to better understand these criteria, we generalize the reflective model creating Figure 4 based on Figure 2.

Where "y_i" is the i^{th} observed variable, "λ_{ij}" is the i^{th} factor loading for the j^{th} latent variable, "η_j" is the j^{th} latent variable, "τ_i" is the intercept for i^{th} observed variable and "e_i" is the disturbance term for the i^{th} observed variable. It is assumed that the disturbance terms have a mean of 0, and are uncorrelated with each other and with η_1.

[6] We think that these criteria are in some sense too strict because no distinction is made between difference in measurement and the cognitive process (Saris & Gallhofer 2007) but in this case we can not make this distinction.

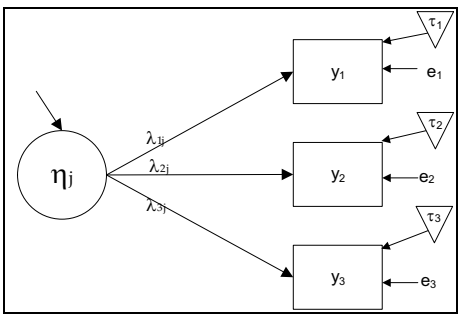
Figure 4. Reflective model with intercepts (τ) and loadings (λ)

$y_1 = \tau_1 + \lambda_{11}\eta_1 + e_1$

$y_2 = \tau_2 + \lambda_{21}\eta_1 + e_2$

$y_3 = \tau_3 + \lambda_{31}\eta_1 + e_3$

$E(e_i) = E(e_i\,\eta_1) = E(e_i\,e_j) = 0 \text{ for } i \neq j$



The analysis for metric invariance is done by testing if the loadings (λ_{ij}) are the same across countries while leaving the intercepts free (metric invariant). If this model holds, it shows that the relationships between this concept and other concepts can be studied, at least if the other concepts are also metric invariant. This test can be done by multiple group analysis (MGA) using any program for structural equation modeling (SEM).

However, metric invariance does not test whether people in a country give systematically higher or lower responses on the response scale. If that is the case the means of the IPIM concept cannot be compared. Therefore a stricter test is needed to show that the means can be compared across countries. This requires that scalar invariance also holds, meaning that also the intercepts in the relationships between the latent variables and the observed variables are the same across countries. So for scalar invariance not only the loadings but also the intercepts have to be the same across countries. This requirement can again be tested with multiple group analysis using a SEM program.

Testing for measurement invariance in ESS1

In order to identify the scale of the latent variable η_1, we constrain one of the loadings (λ_{ij}) to 1 and one of the intercepts (τ_i) to 0. Thus, the other estimates are relative to the mean and scale of the constrained variable.

Normally a test for configural invariance is first done to see if the restrictions of the factor model (shown on the left of Figure 4) hold in all countries. In our case, however, the model with no parameters constrained to be equal across countries is just identified, meaning that no test for the model is possible. We

therefore have to assume that the factor model holds, but we can still test whether its parameters are significantly different across countries.

Testing for metric invariance means that one loading is constrained to 1 ($\lambda_{TV,IPIM}$ in our case) and the remaining loadings are estimated but each one must be equal across countries. Metric invariance does not hold in this multiple group analysis for 21 countries in the ESS first round. Two countries, Italy and Portugal, had loadings that were significantly different from those of all the other countries. This was detected with a procedure which has been developed by Saris, Satorra & Van der Veld (forthcoming). This method determines whether misspecifications are present in the specified model. For more details we refer to the website of the SRC[7]. The result means that in Italy and Portugal people answered these questions significantly differently than in the other 19 countries. If these differences are also substantively relevant their scores cannot be compared with those of the other countries.

Testing for *scalar invariance* in addition to the metric invariance, we constrain the intercepts to be equal across countries. The way the test is performed can be found in the website of the SRC[8]. Unfortunately, it turned out that also in this case several coefficients differed across countries.

The final results of the two tests are shown in Table 2. In the first row of Table 2 the estimates for the 17 scalar invariant countries are mentioned. The other rows present the estimates for the differing countries.

Table 2. Scalar invariance and final estimates for the differing countries in 1st round

	λtv,IPIM	λrd,IPIM	λnwsp,IPIM	τtv	τrd	τnwsp
Scalar Invariance	1	.819	.578	0	.054	.373
Greece	*	*	*	*	-.073	*
Italy	.395	*	*	.286	*	*
Norway	*	*	*	*	.110	.290
Portugal	*	*	.159	*	*	.694

*Invariant estimates.

[7] "Detection of deviating loadings" in http://www.esade.edu/research/src/investigacion/surveydesign.

[8] "Detecting deviating intercepts" in http://www.esade.edu/research/src/investigacion/surveydesign.

In sum, the differences in the response functions can come from the loadings (λ) or from the intercepts (τ). The difference in loadings means that the difference in the observed scores across countries varies with the value on the latent variable IPIM. Differences in intercepts across countries mean that the difference is the same for any value of IPIM.

Figure 5. Graphical scalar invariance and deviations

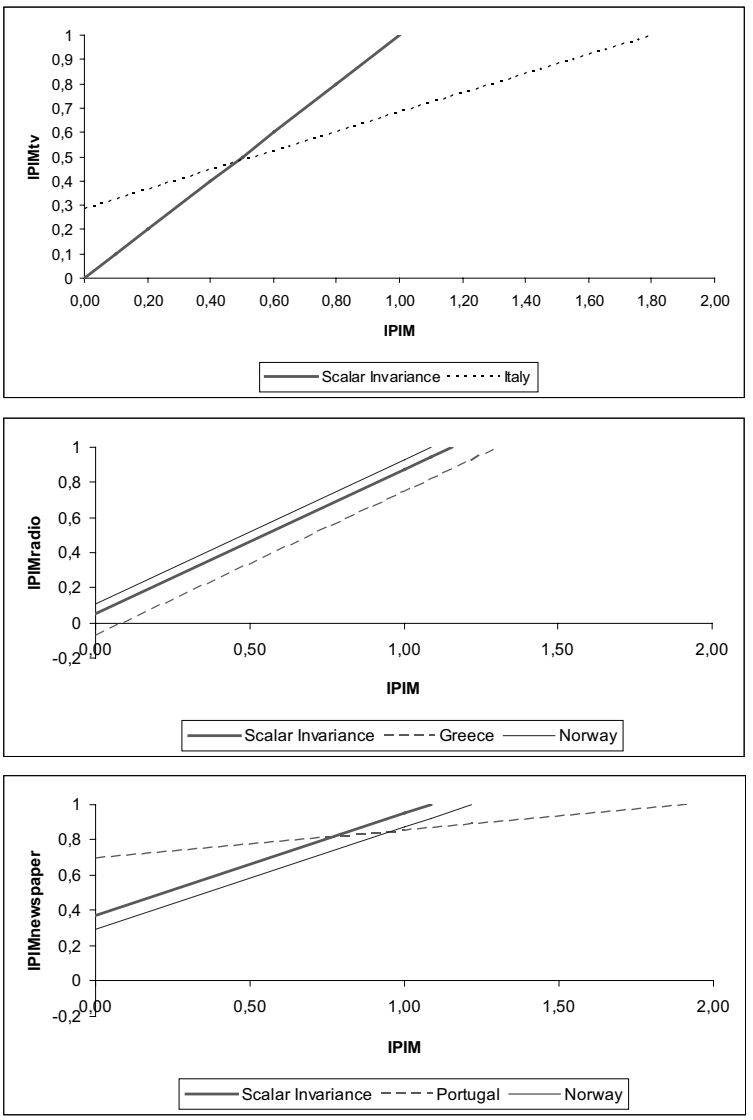


Figure 5 shows graphically the deviations from the scalar invariant relationships presented in Table 2. For $IPIM_{TV}$ the only differing country is Italy, and we can see that it considerably deviates in the loading and intercept from the other countries.

For $IPIM_{Radio}$ the deviations seem not so serious. The two differing countries are Greece and Norway, and they only have deviations in the intercept; therefore the response functions are parallel where the line for Greece is .127 lower and for Norway .056 higher.

For $IPIM_{Newspaper}$ there are two countries differing in scalar invariance. In Portugal the loading and intercept deviate, while for Norway only the intercept deviates; for Portugal we can see that the intercept is larger but the loading is smaller; for Norway only the intercept is .083 smaller.

It has been shown that there are statistically some significantly different coefficients in some countries, however, one can still ask the question: Are these deviations serious enough to create problems in the comparison across countries? In this context we have to realize that comparisons are made for relationships and means and this can be done using the latent variable IPIM or the IPIM index. Saris & Gallhofer (2007) have shown that the most serious problems can be expected if indices are used in these comparisons. Therefore, we will first discuss the construction of the index and then discuss whether the above observed deviations create serious problems in the comparison across countries.

Testing for measurement invariance in ESS2

The same procedure was used for testing measurement invariance in the 2nd ESS round, which is composed of 26 countries[9].

Metric invariance does not hold in this multiple group analysis for all 26 countries. There are two countries, Hungary and Ireland, whose loadings were significantly different from the other countries. This was detected with the misspecification procedure by Saris et al. (forthcoming).

Allowing for these deviations in Hungary and Ireland, we tested again for *scalar invariance*. It turned out that only for these two countries some coefficients differed across countries. The final results of this invariance test for the 2nd ESS round are shown in Table 3. In the first row of Table 3 the estimates for the 24

[9] Austria, Belgium, Switzerland, Czech Republic, Germany, Denmark, Estonia, Spain, Finland, France, Great Britain, Greece, Hungary, Ireland, Iceland, Italy, Luxemburg, The Netherlands, Norway, Poland, Portugal, Sweden, Slovenia, Slovakia, Turkey and Ukraine.

scalar invariant countries are mentioned. The others rows present the estimates for the differing countries.

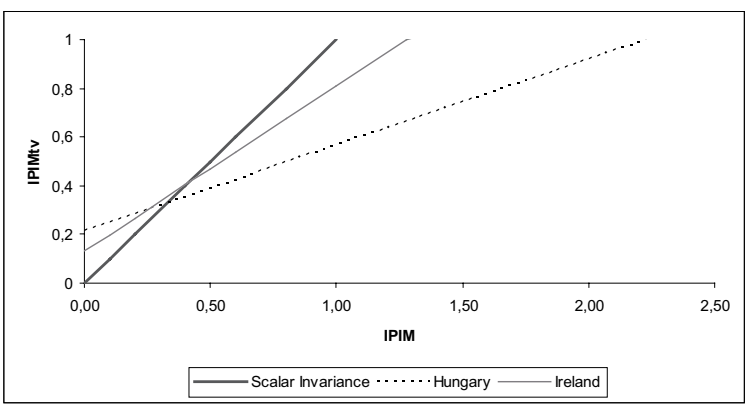
Table 3. Scalar invariance and final estimates for the differing countries in 2nd round

	λtv,IPIM	λrd,IPIM	λnwsp,IPIM	τtv	τrd	τnwsp
Scalar Invariance	1	.800	.563	0	.059	.382
Hungary	.353	*	*	.213	*	*
Ireland	.679	*	*	.131	*	*

*Invariant estimates.

The estimated loading of IPIM on television is very different for Hungary (.353) and Ireland (.679) compared to other countries. This means that the effects are lower for these countries. Also the intercept for television is different. In these countries the curve starts at a higher point (.213 for Hungary and .131 for Ireland) than in the other 24 countries. These differences are shown in Figure 6. It shows the deviations from the scalar invariant relationships presented in Table 3. The deviation exists for $IPIM_{TV}$ in Hungary and Ireland, being more severe in Hungary.

Figure 6. Graphical scalar invariance and deviations on ESS2



Invariance measurement across ESS rounds

In the previous section we measured the metric and scalar invariance for the 1st and 2nd ESS rounds separately. We have seen that some countries are invariant for two rounds, and others are deviant in one round. In order to have a complete picture of the invariance, we also tested for invariance across time for the countries which are in both the 1st and 2nd ESS rounds. It will be carried out for 20 countries, which are Austria, Belgium, Switzerland, Czech Republic, Germany, Denmark, Spain, Finland, France, Great Britain, Greece, Hungary, Ireland, Italy, The Netherlands, Norway, Poland, Portugal, Sweden and Slovenia.

In the specification of the multiple group analysis of this structural equation model, there are 40 groups, which correspond to each country in the two time periods (1st and 2nd ESS rounds). The specification of the metric invariance follows the same approach as we did in the analysis for each round separately, including the deviations found in each round. This means that for metric invariance across time, the loadings (or slopes) were constrained to be equal across countries except for Italy and Portugal in the 1st round and for Hungary and Ireland in the 2nd round.

We tested the hypothesis of the invariance across countries and time for the remaining countries which means that we tested if the relationships between IPIM on television, radio and newspaper are the same across the countries in both ESS rounds.

Using the multiple group analysis we found no more misspecifications for the loadings in the model. Therefore, we can say that the loadings are the same for the remaining 16 countries and for the two time periods.

Testing for scalar invariance is more restrictive because the equality of the intercepts is also needed. The model specified adds the deviations found in the intercepts in the 1st and 2nd rounds (see Tables 2 and 3). The test for scalar invariance detected additional misspecifications for Slovenia in the intercepts on television and radio and for Ireland with respect to the intercept on radio, which were not present at in the single year analysis. This means that there are 15 countries that are scalar invariant within and across rounds. The final estimates are in Table 4.

Table 4. Scalar invariance and final estimates for the differing countries

ESS1=ESS2 Scalar Invariance	λtv,IPIM	λrd,IPIM	λnwsp,IPIM	τtv	τrd	τnwsp
15 countries	1	.911	.559	0	.007	.387
Hungary	.358 (ESS2)	*	*	.206 (ESS2)	*	*
Italy	.390 (ESS1)	*	*	.286 (ESS1)	*	*
Ireland	.684 (ESS2)	*	*	.219 (ESS2)	.168(ESS2)	*
Slovenia	*	*	*	.248 (ESS1)	.179(ESS2)	*
Portugal	*	*	.177 (ESS1)	*	*	.685 (ESS1)

*Invariant estimates.

According to Table 4, there are still 5 deviant countries; Slovenia is metric invariant, therefore, relationships with other variables can still be compared. Hungary, Ireland, Italy and Portugal are not metric invariant, thus for these countries no comparisons can be made with the other countries. The difference in estimates is quite large for these countries. Hungary and Italy are different in slope and intercepts on television as can be seen in Figure 5 for Italy and Figure 6 for Hungary. Ireland is deviant in the slope on television and the intercepts for television and radio, while Slovenia is deviant on these intercepts. Portugal has a distinct pattern, which is different for intercept and slope on newspaper. Thus, the conclusion is that 15 countries are scalar invariant, 16 countries are metric invariant and 4 cannot be compared with the others if the composite scores are used. This does not mean that comparisons cannot be made. The measures for all these countries are partially invariant and therefore the scores of these countries can be compared using the latent variables (Byrne, Shavelson & Muthén 1989). We will come back to this issue later.

The procedure for testing invariance across rounds for each country follows the same approach that we used for the multiple group analysis. In this case, we have two groups; the first group is the country (Italy, for instance) in the first round and the second group is the same country in the 2nd round. Then, we constrained the loadings and intercepts to be equal across rounds.

Following this procedure, we found scalar invariance in Hungary, Italy and Ireland. This means that the interpretation of the paths is the same across time for the individual country. Portugal still has differences across rounds in the loadings and intercepts, which means that this country can neither be compared within nor across rounds using composite scores. In the next section, we will create an index for IPIM and we will see how serious the differences in means are across rounds for these diverging countries, especially Portugal.

The construction of an IPIM index

In this section we discuss the creation of the IPIM index and discuss the quality of it. This composite score is an average of the three media variables. The weights should be identical across all countries, otherwise, the weighting itself will create differences again across countries in the index. So we use the unweighted (weight=1/3) composite score. In this way the composite score for IPIM is the mean over the three components.

Table 5. Comparison of means for 1st ESS round

ESS1 Country	Latent Mean	Composite score mean
Norway	.586	.559*
Portugal	.521	.541*
Switzerland	.513	.504
Sweden	.512	.490
Denmark	.501	.496
Israel	.487	.493
Finland	.481	.470
France	.471	.471
Austria	.463	.468
Netherlands	.457	.452
Ireland	.456	.460
Slovenia	.438	.410
Spain	.424	.445
Belgium	.414	.405
Poland	.414	.405
Germany	.407	.429
Italy	.386	.417*
Czech Republic	.375	.396
Great Britain	.371	.365
Hungary	.360	.378
Greece	.337	.309*

*Not scalar invariant countries.

In addition, the model specified for scalar invariance allows a direct estimate of the means of the latent variable IPIM. It has been shown by Byrne, Shavelson & Muthén (1989) that one can get a consistent estimate of the mean of the latent variable if at least two reflective indicators are scalar invariant. If one would like to compare means of composite scores all indicators on which the scores are based have to be scalar invariance as shown by Saris and Gallhofer (2007).

This is true for only 17 countries in the first round and 24 in the second round. Therefore, we expect on statistical grounds for only 17 countries in the first round and 24 countries in the second round that the means of the composite score will be statistically comparable with the latent variable means.

We have made the comparison between the latent means and the means obtained using the IPIM index developed above. If the deviations in the composite scores are not too severe, these two sets of means should be very similar and correlate highly. Table 5 presents the two sets of means for all countries in the 1st ESS round.

The correlation without all countries for which scalar invariance does not hold in the first round of ESS1 is .958. The correlation between the two sets of estimates - including Greece and Norway with minimal deviations is .965. And the correlation with all 21 countries, including the ones with deviations is .955. These correlations are sufficiently close to 1 to say that the bias in the means must be rather small. This situation is also shown on the scatter plot in Figure 7.

Figure 7. Latent versus composite score mean for the 1st ESS round

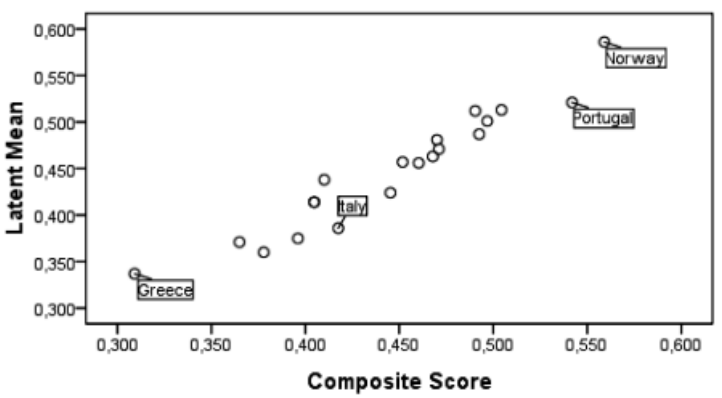


Figure 7 shows that the non scalar invariant countries are not clearly deviating from the others. Even Italy and Portugal, not metric invariant, seem not to be outliers in this plot.

The means can be directly interpreted in the proportion of the time spend on political issues in the media. So we see that in Norway, Switzerland, Sweden and Denmark more than 50% of the time, spent on the media, is directed to political issues. Therefore, the people in these countries are most interested in political issues in the media (television, radio and newspaper). At the other side, Czech

Republic, Great Britain, Hungary and Greece are the countries where the people are the least interested in political issues in the media; they use the media less than 40% of the time for information on political issues.

Table 6 shows the comparison of the latent means versus the composite score means for all countries in the 2nd ESS round including Hungary and Ireland which were not scalar invariant.

Table 6. Comparison of means ESS2

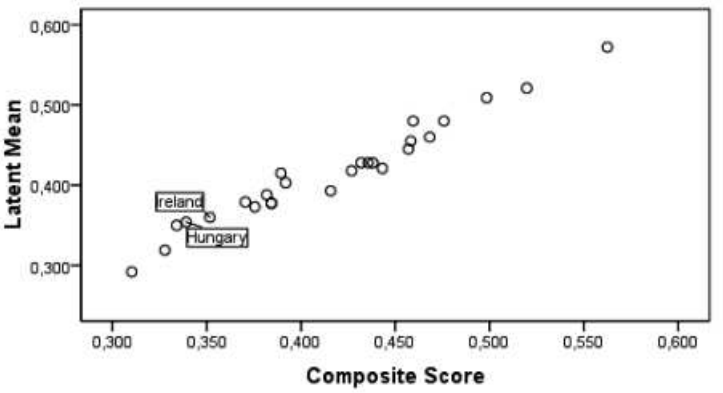
ESS2 Country	Latent Mean	Composite score mean
Norway	.572	.562
Denmark	.521	.520
Sweden	.509	.498
Finland	.480	.476
Iceland	.480	.460
Netherlands	.460	.468
Switzerland	.455	.458
Portugal	.445	457
France	.428	.436
Luxemburg	.428	.438
Ukraine	.428	.432
Spain	.421	.443
Austria	.418	426
Slovenia	.415	389
Poland	.403	.397
Germany	.393	.415
Italy	.388	.381
Slovakia	.379	.370
Turkey	.378	.384
Estonia	.377	384
Belgium	.373	.375
Ireland	.360	.351*
Hungary	.354	.339*
Great Britain	.350	.334
Czech Republic	.319	.327
Greece	.292	.310

*Not scalar invariant countries.

The correlation between the two sets of estimates, calculated over all 26 countries and also over 24 countries (excluding Ireland and Hungary which are not

scalar invariant) is .980, which again is very close to 1. The scatter plot confirms of course this result.

Figure 8. Latent versus composite score mean for the 2nd ESS round



Also in this case the means of the composite scores for the non scalar invariant countries, Ireland and Hungary, seem not to be very different from the latent means.

In this 2nd round, the latent means for IPIM show that again for the Scandinavian countries (Norway, Denmark, Sweden and Finland) more than 50% of the time spent on the media is directed to political issues, following the same pattern as in the 1st round. In comparison, Hungary, Great Britain, Czech Republic and Greece are the countries least interested in political issues in the media; they use the media for less than 40% of the time for information on political issues.

Moreover, the rank ordering of the countries is very similar in the two rounds. The Pearson correlations of the ranking of the means of the latent variables is .90; while the correlation of the composite scores is .875 when the 20 countries were used (also the non invariant countries) and .925 when the 16 invariant countries are used.

The difference between the composite scores within countries, which are scalar invariant across ESS rounds, can be obtained from the difference between values in Table 5 and Table 6.

The bias due to deviations from scalar invariance

In general it is argued that means of composite scores can only be compared if the indicators are scale invariant (Horn et al. 1983; Meredith 1993; Steenkamp & Baumgartner 1998, Saris & Gallhofer 2007). We consider this restriction too strict; we demonstrated in the previous section that the countries which are not invariant, for both rounds, do not have a large deviation in the composite score means. This is so because the differences across countries were mostly found for only one indicator while the composite score is based on several indicators

The bias in the IPIM index due to deviation from scalar invariance can be defined as the sum of the estimated differences multiplied by the weight for the differing media, which means one third for each medium. For instance, in the 1st round for Italy, television was the differing medium. The bias in the mean is calculated to be -.017 because:

$$[(\tau_{\text{scalar invariance}}-\tau_{\text{ITALY}}) + (\lambda_{\text{scalar invariance}}-\lambda_{\text{ITALY}})\bar{\eta}_{ITALY}\ ]w_{TV} =[(.0 -.286) + (1-.395)*.386]* (1/3) =-.017$$

For Portugal the same calculation leads to an estimate of the bias of .03. The estimates of the total biases in the means are small but the bias depends on the value for which the comparison is made. If one wants to compare the means of the composite scores for the means of the latent mean ($\bar{\eta}_i$), which is .386 for Italy and .521 for Portugal, these biases are rather small. The reason is shown in Figure 5, where in the first graph for an IPIM latent mean of .386 the deviation from scalar invariance is quite small, but for other values of the latent variable the bias increases rapidly. This signifies that one could compare in this case the means of the composite scores for these countries with the others as well, but one cannot use the composite scores for other purposes due to the bias.

For the other countries, the calculation of the bias in the composite score is simpler because there is only a deviation in the intercepts. For Norway there is a positive and a negative bias of almost equal size which yields a bias of .01. For Greece the bias is .04. The example of calculation for Greece is as follows:

$$(\tau_{\text{scalar invariance}}-\tau_{\text{GREECE}})w_{\text{Radio}} =(.054 -(-.073)) (1/3) = (.127) (1/3) = .04$$

For the second round, the same procedure is followed and the bias is -.01 for Hungary and -.002 for Ireland. Compared with the mean values this is so small that it will not change the rank ordering of the countries.

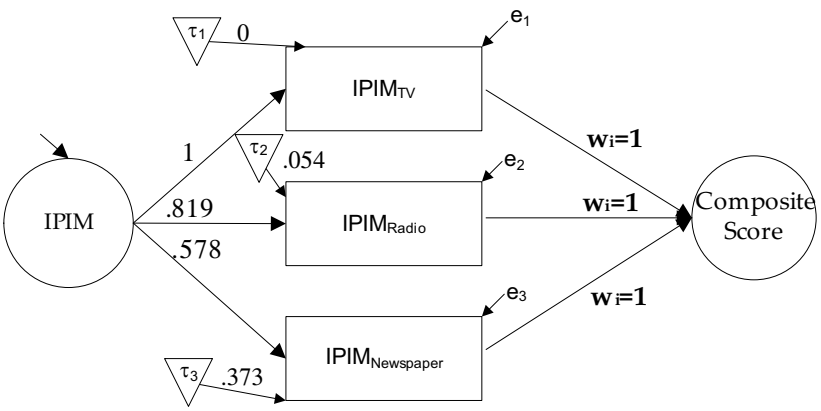
This illustrates that a less strict definition of equivalence of instruments, not only looking at the statistical significance of differences but also at the relevance of the bias in the index, will allow comparison of means more frequently than the more strict definition which is normally used.

Quality of the IPIM index

In the previous section we showed that the composite scores are not seriously biased by a few deviations from scalar invariance in this specific case. Another matter is how good the quality of this index is. This refers to the amount of random error in the index.

The quality for single indicators is defined as the correlation between the theoretical variable of interest and the observed variable. For the composite score this definition can also be used. The model specified for estimation of the composite scores in the 1st round is shown in Figure 9, where loadings and intercepts are from Table 2; loadings and intercepts for the 2nd round are shown in Table 3.

Figure 9. Reflective model and composite score for IPIM in the 1st round



Now we are interested in the size of the correlation between the latent variable of interest IPIM (Interest in Political Issues in the Media) and the Composite Score (S). The quality of the Composite Score as indicator for the latent variable is the correlation squared between these two variables. We can obtain the correlation between these two variables as follows (Saris & Gallhofer 2007):

$$\rho(\text{IPIM, S}) = \frac{\sum q_i w_i}{\sqrt{\text{var}(S)}}$$

Where "q_i" is the i^{th} standardized loading estimated for each country in the multiple group analysis; "w_i" is the weight for the i^{th} indicator which is chosen to be equal to 1/3 for all "i" and "var(S)" is the variance of the index. The variance

of the composite score can be obtained directly with any statistical software or calculated as:

$$var(S) = \Sigma\, w_i^2\, var(var_i) + 2\Sigma\, w_i w_j\, cov(var_i, var_j)$$

where $cov(var_i, var_j)$ is the covariance of the observed variables i and j.

To give an example, for Germany in the 1st round $\rho(IPIM, S)_{DE} = .65$ because $q_1 = .548$; $q_2 = .368$; $q_3 = .243$ are the standardized loadings in Germany for this model[10]. Then, the quality is $\rho^2(IPIM, S)_{DE} = .42$

Table 7. Quality for composite scores

	1st round	2nd round
Austria	.36	.36
Belgium	.30	.35
Switzerland	.32	.27
Czech Republic	.46	.43
Germany	.42	.40
Denmark	.36	.32
Estonia	-	.38
Spain	.31	.33
Finland	.29	.33
France	.27	.28
Great Britain	.34	.37
Greece	.37*	.33
Hungary	.37	.34*
Ireland	.52	.56*
Israel	.37	-
Iceland	-	.36
Italy	.31*	.29
Luxemburg	-	.28
Netherlands	.39	.32
Norway	.27*	.33
Poland	.28	.31
Portugal	.28*	.39
Sweden	.30	.34
Slovenia	.27	.35
Slovakia	-	.36
Turkey	-	.21
Ukraine	-	.36

*Countries with differences.

[10] Note that the standardized loadings are not the same across countries due to the differences in variances of the observed and latent variable.

Using the same procedure for the remaining countries, we obtain the qualities for the composite scores for all countries. The qualities for the composite scores in the 1st and 2nd round are shown in Table 8.

The quality can be interpreted as the percentage of variation in the observed composite score, which can be attributed to the true IPIM variable of interest.

The qualities are rather low because the correlations among the observed variables are also not very high. Given the relative low quality of this index, one can consider improving the index. But given the present situation one should realize that the relationships between the composite score for IPIM and other variables will be considerably underestimated. Therefore it is important to correct for measurement error studying the relationships of the index with other variables. How this can be done has been shown in several papers on structural equation modelling (Bollen 1989; Saris & Gallhofer 2007).

Substantive Relationships of IPIM

Some Expected Relationships

There are several studies which relate the use of the media with political issues. McLeod, Scheufele and Moy (1999) studied the relationship between interest in local politics and use of various types of media. They found a strong relationship between local media use and community integration and local political interest, knowledge, and participation. The study of these authors showed that television and radio news are intervening variables between political interest and political participation (McLeod et al. 1996).

Scheufele (2002) found a relationship between news media use, interpersonal discussion about politics and participatory behaviour. Smith (1986) found a positive and over time constant relationship between newspaper readership and political knowledge. Newspaper readership was itself a major determinant of political knowledge and activity. Furthermore, the relationship remained constant over time among people with similar levels of education and socioeconomic class.

Summarizing other research, several studies found a relation between media use and political participation, for example Newton (1999), Eveland & Scheufele (2000), Aarts & Semetko (2003) or Vreese & Semetko (2004). A relationship between media use and knowledge has also been reported in Guo (2000) and Vreese & Boomgaarden (2006).

This incomplete overview of relevant theory suggests that we should find relationships between media use and political discussion, political participation and political knowledge. The reader should be aware that media use is not the same as "interest for political issues presented in the media" as our IPIM variable represents. However, we suppose that it is reasonable to expect the same relations for our IPIM variable as for the more general media use variable mentioned in the literature. On the other hand, we expect that the IPIM variable will be strongly related as well with the variable political interest.

The model for validation

In this section we develop a model relating political interest, political actions and the composite score for IPIM with political knowledge and political discussion. The model suggested is shown in Figure 10.

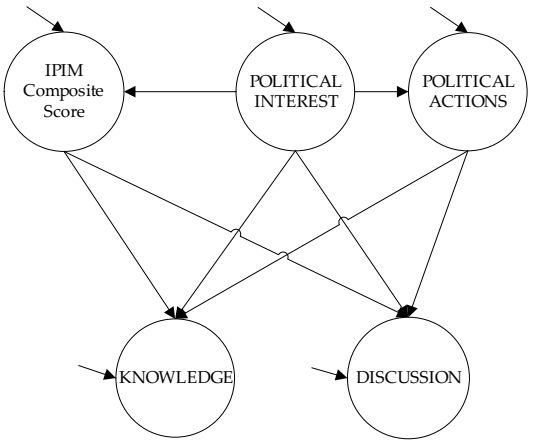
The specification of the model in Figure 10 entails the following hypotheses. The first one is that we expect an influence from 'political interest' on the latent variables 'political discussion', 'political knowledge', and 'political actions'; the reason is that people who are more interested in political issues will discuss politics more often, they will know more about political aspects and they will participate more in different types of political actions.

The second hypothesis is that we expected a strong influence from 'political interest' on the IPIM composite score; the reason is that people who are more interested in political issues in general will also be more interested in political topics in the media.

The third hypothesis suggests a direct and positive effect from IPIM and 'political participation' on 'political knowledge' and 'political discussion'. The interpretation is that people who have a higher IPIM score will participate more in political activities and will discuss more about political issues and will know more about these topics.

However, according the specified model, it may be that 'political interest' causes spurious relationships between several outcome variables, for example between IPIM and discussion, IPIM and knowledge, and IPIM and political actions. Therefore, the fourth hypothesis is that a large part of the correlations between these variables will be spurious, reducing the effect of IPIM on the other variables.

Figure 10. Theoretical model for the relationships of IPIM and political variables



For the measurement of these outcome variables we refer to Appendix 1. For more details on the input for the estimation of IPIM composite score, see the web of the Survey Research Centre[11] of ESADE.

Table 8 shows the correlations between IPIM composite score and the variables political interest (PI), political discussion (DISC), political actions (PA) and political knowledge (KNOW). This table presents in its first column for each country the correlations between the observed variables without correction for measurement error. In order to give an impression of the correlation corrected for measurement error Table 8 contains in its second column for each country the correlations corrected for measurement error. The disattenuated correlation ($\hat{\rho}_d$), corrected for measurement error, between the observed variables was computed as:

$$\hat{\rho}_{d\,ii'} = \frac{\hat{\rho}_{xi\,xi'}}{\sqrt{Q_i Q_{i'}}}$$

where xi and xi´ are the observed variables and Q_i and $Q_{i'}$ are the estimates of the qualities (see Appendix 1) of the variables.

The results show that the IPIM composite score has significant correlations with all different outcome variables in all countries, as expected. This confirms the external validity of the IPIM composite scores. It is also clear that the measurement error has a considerable effect on the correlations. Correcting for it is therefore essential. After correction, the correlations are much bigger.

[11] "Scalar Invariance" in http://www.esade.edu/research/src/investigacion/surveydesign.

Table 8. Correlations between the IPIM Composite score and the variables PI (political interest), Disc (discussion), PA (political action) and KNOW (knowledge)

	Austria		Belgium		Switzerland		Czech Rep.		Germany		Denmark		Spain	
PI	.295	.601	.389	.868	.314	.678	.311	.560	.346	.612	.317	.645	.198	.434
DISC	.252	.640	.246	.685	.256	.690	.269	.605	.265	.624	.163	.414	.187	.512
PA	.186	.372	.210	.460	.194	.411	.160	.283	.248	.449	.164	.328	.085	.183
KNOW	.138	.256	.091	.185	.130	.255	.113	.185	.125	.214	.113	.209	.114	.228

	Finland		France		Great Britain		Greece		Hungary		Ireland		Italy	
PI	.286	.649	.285	.670	.432	.926	.337	.693	.290	.596	.377	.654	.306	.687
DISC	.171	.484	.211	.619	.335	.876	.325	.815	.240	.602	.305	.645	.260	.712
PA	.152	.339	.178	.411	.269	.577	.165	.325	.118	.233	.155	.258	.170	.366
KNOW	.059	.122	.024	.051	.137	.261	.087	.159	.081	.148	.154	.237	.137	.273

	Israel		Netherlands		Norway		Poland		Portugal		Sweden		Slovenia	
PI	.339	.697	.386	.773	.261	.628	.331	.782	.182	.430	.351	.801	.259	.623
DISC	.274	.687	.309	.755	.203	.596	.297	.856	.160	.461	.259	.721	.176	.517
PA	.101	.199	.225	.425	.163	.376	.157	.356	.069	.156	.17	.372	.066	.152
KNOW	.067	.122	.120	.214	.080	.171	.061	.128	.078	.164	.156	.316	.029	.062

Firstly, the parameters of the model in Figure 10 were estimated correcting for measurement errors[12]. The model was corrected for misspecifications using the testing procedure of Saris et al (forthcoming). These corrections for misspecifications has led to the introduction of some effects and correlated disturbance terms which were not specified before[13]. The results for the effects specified in the model of Figure 10 have been presented in Table 9.

[12] Because of the skewed distribution of the variables, the estimation was done with ML and WLS. However the estimates using the two procedures were the same and both procedures indicated the same significant effects.

[13] Significant correlations between the disturbance terms of political knowledge and discussion were found in Denmark (.44), Spain (.76), Greece (.39), Hungary (.93), Ireland (.40), Italy (.46), Poland (.42), Portugal (.39), and Slovenia (.30).

Table 9. Standardized coefficients for the model in the different countries

	IPIM on discuss	IPIM on knowl	PA on discuss	PA on knowl	PI on discuss	PI on knowledge	PI on IPIM	PI on PA
Austria*	.04	.01	-.02	-.03	.98**	.42**	.62**	.50**
Belgium	.01	-.04	.08	.01	.69**	.70**	.88**	.49**
Switzerland	.06	.13	.14*	.06	.87**	.16**	.69**	.52**
Czech Rep*	.04	-.04	.02	.04	.97**	.37**	.57**	.45**
Germany	.05	-.01	.15**	.01	.88**	.32**	.73**	.46**
Denmark	.03	.01	.38*	.04	.68**	.32**	.63**	.41**
Spain	.19**	.11	.46**	.07	.52**	.26**	.42**	.51**
Finland	.03	-.04	.32**	.14**	.51**	.19**	.66**	.46**
France	-.08	-.12	.28**	.01	.83**	.23**	.69**	.55**
Gr Britain*	.01	-.01	.19**	.03	.88**	.28**	.90**	.59**
Greece*	.14**	-.05	.20**	.08	.74**	.27**	.70**	.41**
Hungary*	-.01	-.12	-.02	.05	.99**	.44**	.60**	.31**
Ireland*	.00	.10	.18**	.11**	.86**	.17**	.64**	.47**
Israel*	-.02	.02	.12**	.05	.95**	.13**	.69**	.31**
Italy*	.02	.09	.04	-.04	.90**	.30**	.66**	.65**
Netherlands	.00	-.02	.13**	-.01	.89**	.30**	.79**	.48**
Norway	.01	.08	.30**	.11**	.74**	.07	.65**	.43**
Poland*	.06	.03	.07**	.05	.89**	.35**	.79**	.37**
Portugal	.16**	.03	.18**	-.03	.64**	.39**	.41**	.54**
Sweden*	.09	.11	.23**	.10**	.92**	.20**	.78**	.52**
Slovenia	.02	-.15	.24**	.06	.75**	.36**	.61**	.38**

* The variances of the disturbance term for the latent variable 'political discussion' has been constrained to 0.

These results completely support our first and second hypotheses because we found significant effects from political interest on political actions, discussion, knowledge, and IPIM composite score for all countries.

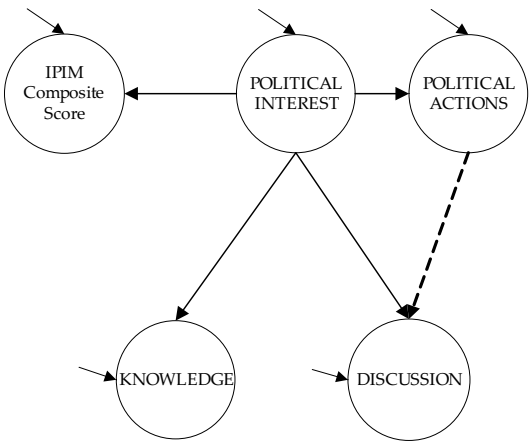
Our third hypothesis is not supported because non significant effects were found for IPIM on knowledge and from IPIM on discussion (except for Spain, Greece and Portugal). This is due to the fact that, in line with the fourth hypothesis, political interest can explain the relationships between the IPIM variables and others variables as spurious relationships. It is even so that the significant

correlation between IPIM and the variables political discussion and knowledge is completely spurious due to the general political interest.

Political actions have also no effect on political knowledge but there are significant effects on political discussion for 16 out of 21 countries. This means that participating in political actions leads to more discussion about these topics which seems to be a logical effect. Additionally, as we expected, we did not found any effect from IPIM on political actions.

Figure 11 shows the final model indicating the significant direct effects that we found. A full line in the figure means that there are significant direct effects in all countries, and the dashed line indicates that in some of the countries the effect is non significant.

Figure 11. Final empirical model for most of ESS countries



The strongest effects from political interest, according Table 10, are on the IPIM composite score and on discussion. This means that people who are more interested in politics spend more time for political issues in the media, participate more in actions, tend to discuss more about these issues and know more about political topics.

We also find that interest in political issues in the media (IPIM) is correlated with political participation, political discussion and political knowledge as expected on the basis of the literature, but in this study we find that these relationships are spurious relationships due to the strong effects of political interest on these other political variables. This is a new observation which does not contradict the previous finding but adds more precise knowledge to what already was known.

Conclusions

In this paper we have studied the operationalization and the quality of the composite scores for the concept "Interest in Political Issues presented in the Media" briefly denoted as IPIM. The purpose of this study was to provide to the ESS data users scores for this construct including information about the quality of the scores which indicates what these scores can be used for and what not. This effort was made so that researchers could avoid having to go through the difficult and sometimes rather technical steps to reach the above mentioned goal and be able to use the scores for the variable IPIM in their research. The scores for all people in all countries between on and three IPIM variables separately and on the IPIM composite score can be found in the ESADE website[14].

Our measure for interest in the media is a combination of the ratio of use of the different media for information about politics and currect affairs. One may make the argument that people could also be highly interested in politics but not necessarily obtain their information from all three sources; maybe they are using only one medium for this purpose. In that case, we could better have created a measure for political interest based on that one source. Computationally that could be done by using the highest ratio for one of the three media as the final measure for IPIM.

In order to check this option we tested with this measure the model in Figure 10. The estimates with this new measure were compared with the results presented in Table 9. It turned out that this analysis did not lead to any improvement of the results. In contrast less significant effects and correlations have been found[15]. Given these results we have concluded that the index based on the combination of the the three media is better. Below we will summarize the results which have been obtained with this composite score. Our analyses in this paper have shown the following possibilities for the use of these scores.

Comparison of means

In the first round of ESS we detected only for Italy and Portugal serious deviations from metric invariance and minor deviations from scalar invariance for

[14] "IPIM Composite Score" in http://www.esade.edu/research/src/investigacion/surveydesign.

[15] This may be partially due to the fact that we could not correct for measurement error in this analysis because we had no estimate of data quality for this new measure.

Norway and Greece. In the second ESS round only two countries, Ireland and Hungary), did not satisfy the criteria for scalar invariance. The literature suggests that the means of the composite scores of non scalar invariant variables can not be compared across countries. However we computed the composite scores for all countries and looked at the bias due to the deviations in the means of the composite scores. It turned out that these biases were so minimal that we think that the means of the composite scores can be compared across countries in round 1 and round 2 for all countries.

That does not mean that one can use the composite scores for Portugal and Italy in round one for other comparisons (for other values of the latent variables for example). This is clearly impossible due to the lack of metric invariance.

Comparison of relationships

Another aspect of the evaluation concerns the size of the random measurement error. This is evaluated by the correlation between the latent variable IPIM and IPIM composite score. It turned out that this correlation was rather low, which suggested that better measurement would be needed and that, for the time being, correction for measurement error is essential in the analysis of the relationship between IPIM composite scores and other variables.

Given the lack of metric invariance for the countries Italy and Portugal the composite scores can not be used to study the relationships with other variables but the latent variables can still be used.

The literature suggested that significant relationships were to be expected between IPIM and political participation, political discussion and political knowledge, and we added to this the relationship with political interest. These correlations were indeed found in the different countries. This is an indication of the validity of the composite score for the IPIM variable.

Appendix 1: Measurement of variables, testing for external validity

Political interest (polintr): Political interest is measured in the ESS as: "*How interested would you say you are in politics?"*. The answers to this question had to be expressed on a 4 point scale from "very interested" to "not at all interested". The quality for this question has been estimated by SQP and turned out to be .76.

I***nterpersonal discussion about politics (discpol):*** This variable is measured by the frequency of political discussions: "*How often would you say you discuss politics and current affairs?"*, it uses a 7 point scale from "every day" to "never". The quality of this measure is estimated by SQP and turned out to be .43.

Political knowledge*:* In the ESS main questionnaire there is not a direct question about the level of political knowledge. So we defined a new variable with several questions available in the 1st, 2nd and 3rd ESS rounds for further comparisons. Political knowledge is defined as the number of times the respondents answered "don't know" on questions concerning ten political issues. Then the value was reversed in order to obtain a value of political knowledge. This means that people who answer less times "I don't know" to these questions are seen as having more political knowledge. The quality of this measure is quite low; it turned out to be .26. The variables used are shown in Table A1:

Table A1. Questions for "political knowledge"

Variable	Label ESS database
Placement on left right scale	lrscale
Satisfaction with present state of economy in country	stfeco
Satisfaction with the national government	stfgov
Satisfaction with the way democracy works in country	stfdem
State of education in country nowadays	stfedu
State of health services in country nowadays	stfhlth
Government should reduce differences in income levels	gincdif
Gays and lesbians free to live life as they whish	freehms
Ban political parties that wish overthrow democracy	prtyban
Modern science can be relied on to solve environmental problems	scnsenv

Political participation*:* In the ESS political participation is operationalized by the unweighted sum of two formative indicators namely *number of conventional actions done* and *number of protest actions done.* In the ESS, the *number of conventional actions done* is measured by "contacted a politician, government or local government official?", "worn or displayed a campaign badge/sticker?" and "donated money to a political organisation or group". *The number of protest actions done* is measured by whether or not the respondents "signed a petition", "Taken part in a lawful public demonstration", "boycotted certain products", "deliberately bought certain products for political, ethical or environmental reasons" and "participated in illegal protest activities".

The quality of conventional and protest actions are calculated using the

unweighted summated scores. Guillén et al., (forthcoming) computed the quality for conventional and protest actions for Germany (.63; .67), England (.61; .64) and The Netherlands (.53; .65) using Survey Quality Predictor (SQP). SPQ is not available for other languages and therefore we assume the quality of the two participation variables to be equal to the average of the three countries (.59 for conventional actions and .65 for protest actions). Based on these values also the quality of the combined index for political participation was calculated which was .73 for Germany, .64 for Great Britain and .72 for The Netherlands.

References

Aarts, Kees and Holli A. Semetko. 2003. "The Divided Electorate: Media Use and Political Involvement." *The Journal of Politics* 65(3): 759-784.

Bollen, Kenneth A. 1989. *Structural Equations with Latent Variables*. New York: Chichester, Wiley.

Byrne, Barbara M., Richard J. Shavelson, and Bengt Muthén. 1989. "Testing for the Equivalence of Factor Covariance and Mean Structures: The Issue of Partial Measurement Invariance." *Psychological Bulletin* 105(3): 456-466.

Coromina, Lluís, and Willem E. Saris (forthcoming). Quality of the Measurement of Media Use on Political Issues in the ESS.

Eveland, William P., and Dietram Scheufele. 2000. "Connecting News Media Use with Gaps in Knowledge and Participation." *Political Communication* 17: 215-237.

Guillen, Laura, Lluís Coromina, and Willem E. Saris, (forthcoming). Measurement of Social Participation and its Place in Social Capital Theory.

Guo, Zhongshi 2000. "Media use Habits, Audience Expectations and Media Effects in Hong Kong's First Legislative Council Election." *Gazzette* 62(2): 133-151.

Harkness, Janet, Fons van de Vijver, and Peter Mohler. 2007. *Cross-Cultural Survey Methods*. New York: Wiley.

Horn, J. L., J. McArdle, and R Mason. 1983. "When is Invariance not Invariant: A practical Scientist's Look at the Ethereal Concept of Factor Invariance." *The Southern Psychologist* 1: 179-188.

Litte, Roderick J. A., and D.B. Rubin. 1987. *Statistical Analysis with Missing Data*. New York: Wiley.

Lynn, Peter, Sabine Häder, Siegfried Gabler, and Seppo Laaksone. 2007. "Methods for Achieving Equivalence of Samples in Cross-National Surveys." *Journal of Official Statistics* 23(1).

McLeod, Jack. M., Katie Daily, Zhongshi Guo, Eveland, William P., Jan Bayer, Seungchan Yang, and Hsu Wang. 1996. "Community Integration, Local Media use, and Democratic Process." *Communication Research* 23: 179-209.

McLeod, Jack M., Dietram A. Scheufele, & Patricia Moy. 1999. "Community, Communication and Participation: The Role of Mass Media and Interpersonal Discussion in Local Political Participation." *Political Communication* 16: 315-336.

Meredith, William. 1993. "Measurement Invariance, Factor Analysis and Factorial Invariance." *Psychometrika* 58: 525-543.

Newton, Kenneth. 1999. "Mass Media Effects: Mobilization or Media Malaise?" *B. J. Pol.* 29: 577-599.

Newton, Kenneth. 2000. European Social Survey (ESS) report on questionnaire development. Media and communication question. Chapter 4, 122-175. Retrieved on 15th January 2007: http://www.europeansocialsurvey.org/index.php?option=com_docman&task=doc_view&gid=123&Itemid=80

Oberski, Daniel and Willem E. Saris. 2007. *Differences in Quality of Questions Across Countries. Festschrift for Jaak Billiet.* University of Leuven.

Saris, Willem E. and Daniel Oberski (forthcoming). "Categorisation Errors and Differences in the Quality of Questions Across Countries."

Saris, Willem E. and Irmtraud N Gallhofer. 2007. *Design, Evaluation, and Analysis of Questionnaires for Survey Research*. Wiley.

Saris, Willem E., Satorra, Albert, and William van der Veld (forthcoming). Testing Structural Equation Model or Detections of Misspecifications?

Saris, Willem E., Albert Satorra, and Dag Sörbom. 1987. "Detection and Correction of Structural Equation Models." *Sociological Methodology* 17: 105-131.

Scheufele, Dietram A. 2002. "Examining Differential Gains from Mass Media and Their Implications for Participatory Behaviour." *Communication Research* 29: 46-65.

Smith, Hampden H. 1986. "Newspaper Readership as a Determinant of political Knowledge and Activity." *Newspaper Research Journal* 7: 47-54.

Steenkamp, Jan-Benedict E. M. and Hans Baumgartner. 1998. "Assessing measurement Invariance in Cross-National Consumer Research." *Journal of Consumer Research* 25: 78-90.

Little, Roderick J. A. and David B. Rubin. 1987. *Statistical Analysis with Missing Data*. New York: John Wiley.

Vreese, Claes H.de., and Boomgaarden, Hajo (2006). "News, Political Knowledge and Participation: The Differential Effects of News Exposure on Political Knowledge." *Acta Politica* 41: 317-341.

Vreese, Claes H.de. and Holli A. Semetko. 2004. "News matters: Influences on the vote in the Danish 2000 euro referendum campaign." *European Journal of Communication* 43: 699-722.