UNIVERSITAT POMPEU FABRA

RECSM

Research and Expertise Centre for Survey Methodology

RECSM Working Paper Number 25

2012

Universitat Pompeu Fabra - Research and Expertise Centre for Survey Methodology
Edifici ESCI-Born - Despatxos 19.517-19.527
Passeig Pujades, 1 - 08003 Barcelona
Tel.: +34 93 542 11 63 web: http://www.upf.edu/survey/ e-mail: recsm@upf.edu

ESS-DACE

The European Social Survey - Data for a Changing Europe

Contract Number: 262208

Deliverable 12.1

Evaluation of questions and concepts - report 1 (Media use)

Start date of project: July 2010 Duration: 48 months

Organisation name of lead contractor for this deliverable: UPF

Dissemination level: PU (PUBLIC)

Due date of deliverable: June 2011

Actual submission date: October 2011

Measurement of concepts based on the media module of the ESS

Lluís Coromina[1]

Willem Saris

Laur Lilleoja

Abstract

Given the importance of the media in all societies, the Central Coordinating Team of the European Social Survey (ESS) asked Ken Newton to propose a module for Media use which could be used in the core questionnaire of the ESS. Together with the Central Coordinating Team of the ESS, a module for media use has been developed. This module allows the measurement of the "total time spent on the traditional media" (TV, Radio and Newspapers), "the total time spent on political issues and current affairs in the media" and "the total time spent for other purposes in the media". Besides that one can measure the "Interest of people for political issues in the media". These measures are evaluated in this report. The questions asked are:

- How are these concepts operationalized?
- Can these measures be compared across countries?
- How good are these measures and do we need these aggregated variables or should we rely on the separate measures of the use of different media?
- How should we compute optimal composite scores for these concepts?
- How can we analyze the relationships of these variables with other variables taking into account the measurement errors in these variables

For the concepts which have been evaluated in a positive way " Use of the media for political issues", "the use of the media for entertainment" and "the interest in political issues in the media the concepts" the scores of all respondents for these concepts have been computed and placed in a data file. This data file can be added to the ESS data files and can be used for further analysis[i].

[1] Respectively from the Universitat de Girona, Universitat de Pompeu Fabra and Tallinn University

Introduction

Given the importance of the media in all societies, the Central Coordinating Team of the European Social Survey (ESS) asked Ken Newton to propose a module for Media use which could be used in the core questionnaire of the ESS. In his paper for the ESS, Newton (2000) suggested that there are three main topics for a questionnaire on media use:

a) Amount of media use. The aim was to ask respondents how often they use the different types of media.

b) The purpose of its use. The media can be used because one is interested in politics, sports, culture, movies etc. So it is relevant to know for what purpose people use the media.

c) Sources of information. It is also interesting to know which type of media is the main source of information for different topics.

The first decision in the design of the media use module was to clarify the type of media (traditional and/or new media) that should be taken into account. "Traditional media" mainly refers to television, radio and newspapers. In contrast, new forms of communication are primarily related to Internet, e-mail, websites of information and related communications. If both are used, do we treat traditional and new media equally or examine them separately? Newton commented on this issue:

> "If television is driving out radio and the printed word, and if the newest digital multi-media communications are likely to further diminish the importance of newspapers, radio… then there is little point in wasting time on the old media. If, in a few years time, most people are surfing the web for information about buying a house or the weather forecast, then local radio and newspapers will have a very low impact. And if a large proportion of the population is on the web for business and private reasons, then advertising money is likely to follow them, so bringing about even more rapid change. The political impact of email and the web may be just as large". (Newton, 2000)

Given these comments it is a serious question whether the ESS should shift from traditional media to new media? However, one should decide on the basis of empirical data how fast these changes occur. Looking at the data from Eurobarometer 1999 and 2006 we get the picture presented in Table 1. In 1999, 71% of the respondents in the 15 countries in the European Union (EU) still looked at the news on the TV and 41% listened to news on the radio and followed the news in the newspaper on a daily basis.

	TV News *Daily*	Read Paper *Daily*	Radio News *Daily*
EU15, 1999	71%	41%	41%
EU25, Spring 2006	66%	35%	42%

Source: Eurobarometer

Table 1: Use of the news media in European Union countries

According to Newton's hypothesis, the percentage of daily time spent on news should decrease significantly from 1999 to 2006. Table 1 shows that the frequency has only slightly decreased: television and radio use for daily news decreased around 5%, and radio even increased 1% from 1999 to 2006. Since the old media is in general still heavily used by people, ignoring the traditional media would produce an important bias in the measurement of the media use. However, in a few years time any mass media survey should cover the old media alongside the new ones.

The next question to be answered when one would like to develop a module on media use concerns the purpose for which people use the media. Newton commented on this issue:

> "Most forms of the mass media may be used for a wide variety of different purposes, not only for amount of media use itself. In fact few are like the cinema that has only one dominant purpose - entertainment. The rest – certainly newspapers, TV, radio, and the web - cover almost every area of human activity - news and current affairs, sport, leisure, arts, education, information, science, business, entertainment, gossip about public figures, and so on". (Newton, 2000)

In order to obtain the complete information about the media use of the respondents, one should know the types of programs people spend time on. Complex questionnaires like diaries have been developed especially to collect this information. However, such measures were not possible in the context of a general-purpose questionnaire, as the ESS aspired to be. There was limited space for the media use module.

Therefore, an effort was made by Newton and the ESS specialists to find a combination of the different aims of the module in one single approach. In Figure 1 you find a list of possible purposes for use of the media. Figure 2 presents a first suggestion to an integrated approach. The idea was to ask the people to specify how much time they spend for each purpose in each medium on a normal day. It would have been possible to ask about the last week or yesterday. The former approach is hampered by memory effects. The latter is rather incidental. Asking for a normal day is not a simple question but it seems a good compromise.

Media can be used for different purposes, see the Media card. Can you estimate *for how* many minutes you use the *(TV*, Radio…) *on a normal day for these* different purposes*?*

Media Card: Different purposes of use of the media:

- entertainment	= quizzes, lotteries, games, shows, …
- politics	= news, actuality, political discussions
- business	= financial information, business information
- sport	= reports about sport events or previews
- hobbies	= gardening , home improvement, panting, holidays, …
- education	= educational programs, science and technology
- arts	= movies, music , discussions about it

Transformation of hours in minutes
1 hour = 60 minutes → 8 hours = 480 minutes

Figure 1: Proposal of design for media use and purpose of its use answers

	Entertainment	Politics	Business	Sport	Hobbies	Education	The arts
Television							
Radio							
Newspaper							
Magazine							
Internet							
E-mail							
Teletext							

Figure 2: Response matrix for media use and purpose of its use

This approach is attractive because adding up the answers in a row one would get the media use of a person for a specific type of medium. For instance, summing the time spent for each purpose for television, one can easily obtain the total time spent on television and this will be correct if all purposes are asked. The ratio of the amount in a cell in each row over the total of the row, gives the relative importance of the different purposes of use for each medium. This information could be used for trend analysis and for analysis at an individual level.

Another option is to add the answers in the columns and the result is the use of the media for different purposes at an individual level. This also can be used for trend analysis and individual analysis. With this straight computation in the matrix structure one can obtain the source of information for each topic. For instance, the cells of the column for politics can be compared with the other topics in order to know which type of media they use for the different purposes. For these comparisons one can use the absolute values or the proportion of the time spent on the media for different purposes.

Although the approach is indeed very attractive, the matrix in Figure 2 requires a lot of time and may result in large errors (Slater 2004) because the respondents must give an estimate of time for each cell, which means 49 answers (7 rows x 7 columns) have to be given. In order to reduce the time, a simplification was suggested preserving the most important components of the approach. The final version of the ESS media use module was based on six questions. These questions are about the total time spent and the time spent on political issues and current affairs on the most used traditional media (television, radio and newspaper). Subtracting both approaches one can get the time spent in other issues in the traditional media. The focus is on political issues and current affairs because these variables can be related with several other ones in the ESS survey. The definitive questions for the media module in the ESS questionnaire for round 1 to round 4 are shown in Figure 3.

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 do not watch 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 programmes 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

Figure 3: Questions for the media use module

It will be clear from Figure 3 that the ESS asks for exposure time. This concept should not be confused with attention, comprehension or retention (Slater 2004). The information we obtain is just an estimate of the time that the respondent spent watching television, listening to the radio or reading the news.

In the ESS media use module there is also a question requesting the total time spent on the new media (internet). However, there is no request on the total time

spent on politics and current affairs for that medium. Due to the lack of comparability with respect to the measurement of the new media use with the measurement of the use of the traditional media, we decided not to include this question in the operationalization of the different concepts for media use we are interested in; *Total Time spent on Political Issues in the Media* (TTPIM), *Total Time spent on Entertainment in the Media* (TTEM) and *Total Time spent in the Media* (TTM). This means that the measures are restricted to the use of the traditional media.

Another possibility of the media module is to measure the Interest in Political issues presented in the media. We denote this concept with IPIM. This concept is defined as the difference in attention people spend on political issues in the media.

In part 1 of this report we pay attention to the measurement of Total time spend at the media i.e. the three concepts TTPIM, TTEM and TTM while in part 2 we will discuss the quality of the measurement of IPIM. Finally in part 3 we will discuss suggestions for improvement of the measures based on the new developments in the media world.

Because of the fact that all concepts we want to evaluate require numeric calculations we start with an adjustment of the used scale for this purpose.

1 Selection of the response scale

One of the problems in the data collection by questionnaires is the selection of the response scale that has to be used. Depending on this choice, the quality of the data can vary significantly. Therefore, in an ESS pilot study a MultiTrait MultiMethod experiment (Campbell & Fiske 1959) was done in order to estimate the validity and reliability of the possible alternatives. The purpose of the experiment was to select the best method for the core questionnaire of future ESS rounds. The MultiTrait MultiMethod (MTMM) design uses at least three traits, each measured with three different methods (Andrews 1984; Saris & Andrews 1991; Saris & Gallhofer 2007a). For these experiments the Split Ballot MTMM (SB-MTMM) design (Saris et al. 2004) was used. It is an alternative to the classical MTMM design (Andrews 1984).

Results from the pilot study in The Netherlands and United Kingdom were described by Saris & Gallhofer (2007b). They remarked that while the validities were approximately the same and high (.95 and higher) for all three methods there were considerable differences in the reliabilities. For all three traits the reliability of the "8 numerical categories" was the highest for both countries. The "numeric open question" method, which was based on hours and minutes, was clearly the worst for all three traits in both countries. Therefore the "8 numerical categories" method was used in the next ESS rounds.

The three methods (response scales) are shown in Appendix 1. The traits used are the questions from Figure 3 about the total amount of time spent in television (A1), radio (A3) and newspaper (A5). So in the first round the scale with "8 numerical categories" was used in the main questionnaire but in a supplementary questionnaire (filled in after the main questionnaire) a "numeric open question" and a scale with "7 verbal categories" were used.

This experiment was repeated in the main data collection in round 1 in all countries in order to get an estimate of the data quality for these questions in all countries of the ESS. For round 2, 3 and 4 this SB-MTMM design was not repeated on the media variables; therefore estimates of the data quality for the countries which were not in the first ESS round are not possible to obtain. We will focus on this point in the next sections.

An analysis[2] of the SB-MTMM for Austria is presented and the results are shown in Table 2. For more details, model input can be found on the web of the Research and Expertise Centre for Survey Methodology (RECSM)[3].

Austria	Reliability			Validity		
	8 num cat	Num open	7 verbal cat	8 num cat	Num open	7 verbal cat
TTM_{TV}	.96	.74	.71	1.00	.96	.87
TTM_{Radio}	.85	.76	.77	1.00	.98	.92
TTM_{Newsp}	.76	.71	.66	1.00	.87	.81

Table 2: Reliabilities and validities for Austria

In Table 2 we can clearly see that the method "8 numerical categories" is the most reliable for all three traits. The reliability is the standardized coefficient squared from the SB-MTMM model, the validity is validity coefficient squared. The "7 verbal categories" is clearly the least reliable and valid, due to its vague indicators. According to these results and the previous ones by Saris & Gallhofer (2007b), the "8 numerical categories" is the most adequate method; therefore this scale has also been used in the core questionnaires for next ESS rounds. However, the "8 numerical categories" method has the problem of not being numerical in the sense that one cannot make numeric calculations with the responses. For this reason, we transformed the categories of this scale to numerical values. The procedure we followed is explained below.

[2] Results have been obtained after data transformation explained in the next subsection.
[3] "SB-MTMM Input" in *http://www.upf.edu/survey/working/extrafiles*

1.1 Transformation of the categories

Each category (except the 8th category) is originally labelled with time reference points. The middle of the reference points is used for numerical transformation for all categories except the last one. For the last category, there is no upper reference point and it is uncertain 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 were asked directly for a numeric estimate of the amount of time spent on each medium[4]. Using these variables a value is derived for the last category in each country. Details of the whole procedure are presented in Appendix 2 (section 2.1). The results were, for instance, that the numerical values obtained in Austria for the 8th category are 234 minutes for television, 271 for radio and 180 for newspaper. For Spain they are 261, 300 and 257 or for Belgium they are 269, 300 and 300, respectively.

Table 3 presents the transformation from original scores into minutes. The newly created numeric values of these variables can be found on the website of the Research and Expertise Centre for Survey Methodology[5].

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

Table 3: Original and transformed media use measures

The values obtained in each country are used also in the 2nd, 3rd and 4th ESS rounds. However, the participating countries in these ESS rounds were not exactly the same as the ones in the 1st round. For instance, Estonia was participating in the 2nd, 3rd and 4th rounds, Russia only in the 3rd and 4th round; consequently, additional information about these new countries was missing. But we had to transform the categorical variables in numerical variables also for these countries. To solve this problem we specified for the higher category for each new country the amount of

[4] Data can be obtained from the website *http://www.europeansocialsurvey.org* or *http://ess.nsd.uib.no*

[5] "Scores" in *http://www.upf.edu/survey/working/extrafiles*

minutes of a country from the 1st round, which has the most similar distribution for each medium. The procedure is explained in the Appendix 2 (section 2.2). We found that for countries not present in the 1st round, we could use in most cases data from a neighbouring country. This means that the differences on media use between countries which are close together geographically are smaller those than between countries that are more distant from each other.

Now that the scale has been defined we can start with the evaluation of the concepts. We start in part 1 with the evaluation of the measurements of time use. In part 2 we will evaluate the concept Interest in political issues in the media

Part 1 The measurement of Total time spend at the media[ii]

The concept by postulation (Saris & Gallhofer 2007a) -construct- *Total Time spent on Political Issues in the Media (TTPIM)* is defined as the total amount of time spent on politics and current affairs in the media: television (A2 in Figure 3), radio (A4) and newspapers (A6). These observed variables are formative indicators for TTPIM; this is because the total time spent on political issues in the media is defined as the unweighted sum of time spent on politics in television ($TTPIM_{TV}$), radio ($TTPIM_{Radio}$) and newspaper ($TTPIM_{Newspaper}$). The formative nature of a model supposes that all indicators that compose the construct should be included in the measure. However, we are only using these three traditional media because these are the ones available in the ESS data. So adding up these values we are able to compute the time spent on politics and current affairs in the traditional media (TTPIM) for a person.

The concept by postulation, *Total Time on Entertainment in the Media (TTEM),* is the difference between TTM and TTPIM for each medium. The formative structure for TTPIM is also used for TTEM, which will be computed as the sum of the differences in answers to the questions A1-A2, A3-A4 and A5-A6 mentioned in Figure 3. It should be clear that we call this behaviour "entertainment" even though it is a summary of all the purposes in Figure 2 except "politics". Hereafter, our measures of so-called "entertainment" are the differences between total time and time spent on political issues.

The last concept by postulation is *Total Time spent in Media use (TTM)* which is defined as the total amount of time spent watching television (A1 in Figure 3), listening to the radio (A3) and reading newspapers (A5). This construct has also a formative structure and is also limited to the traditional media.

2 Operationalization of the constructs Total Time spent on the media

Each of these three concepts by postulation (TTPIM, TTEM and TTM) is defined as sum scores of three observed variables (indicators) measuring the time spent (exposure) watching television, listening to the radio and reading newspapers for different reasons.

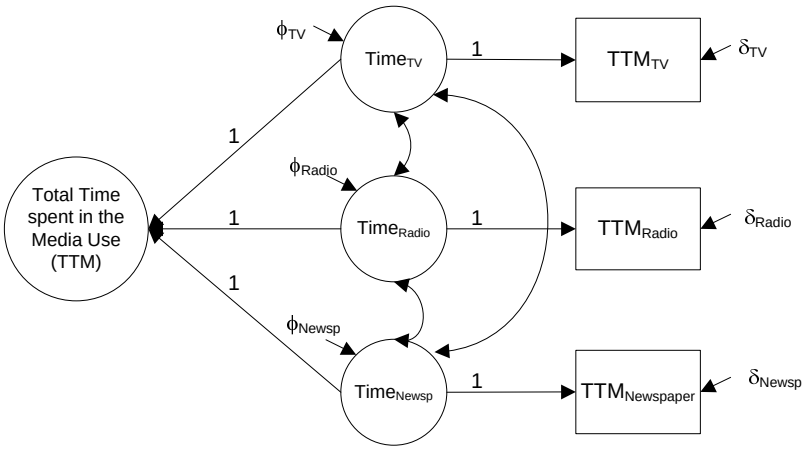

Figure 4: TTM Formative model

Since the models of the operationalization of the three concepts by postulation are the same, we will present only one model, especially the model for the total time spent on the media (TTM) as an example for the three models. Figure 4 presents this model.

The squares in Figure 4 are used to indicate the observed variables (indicators), which are the answers to the questions A1, A3 and A5 in Figure 3 for TTM. In the same way the answers to the questions A2, A4 and A6 can be used as indicators for TTPIM and the answers to the questions A1 minus A2, A3 minus A4 and A5 minus A6 used for TTEM. The answers certainly contained errors, which are indicated as δ_i. Therefore, each observed response variable is connected with the variable representing the concept by intuition, which are concepts for which the measurement is obvious (Saris & Gallhofer, 2007a). The names of these variables are presented in a circle, being unobserved (latent) variables with variances and covariances denoted as ϕ_{ij}. The latent variable "Total Time spent in the Media Use (TTM)" is the behavioural variable of interest. This concept by postulation (construct) is defined as the sum of the observed responses over all three media, corrected for measurement error for each respondent. The model allows for correlation between the use of the different media.

The model in Figure 4 can be represented in the following equations:

$TTM_i = Time_i + \delta_i$ for i= TV, Radio, Newspaper

$TTM = Time_{TV} + Time_{Radio} + Time_{Newspaper}$ (1)

$E(\delta_i) = 0$ for i= TV, Radio, Newspaper

$Cov(TTM, \delta_i) = 0$ for all i

$Cov(Time_i, \delta_j) = 0$ for all ij

$Cov(\delta_i, \delta_j) = 0$ when ij

The last three equations indicate that there is neither covariance between the error variables and the endogenous latent variable (TTM) and exogenous latent variable ($Time_i$) nor between the error variables.

The model specified in Figure 4 cannot be estimated because it is not identified. Formative models specify that the observed variables (indicators) are the cause of the concept by postulation (latent variable); this means that the concept by postulation is a latent dependent variable so that by definition the model is not identified. However, adding additional observed variables influenced by the latent variable will solve the underidentification problem. For this reason, we will use a MIMIC (Multiple Indicators Multiple Causes) model to evaluate whether the definition of the concepts by postulation total time spent on the media make sense. We will come back to this issue in a later section. In the next sections, we will discuss the quality of the composite scores computed for the concepts by postulation.

3 Comparability of the composite scores

In the model of Figure 4 the variable TTM is defined as the sum of three latent variables. Differently, the composite score can, of course, only be calculated using observed scores. Thus, the calculation of the composite scores for the total time spent in the different media variables is done as follows:

$$CS\_TTM = \sum_{i=1}^{3} TTM_i = \sum_{i=1}^{3} \left(Time_i + \delta_i \right) \quad \text{for i= TV, Radio, Newspaper} \qquad (2)$$

Given that the indicators contain an error component, the composite scores will also contain errors and therefore we will evaluate the quality of the composite scores for the different concepts by postulation in the next section. This comparison will be done for the three different ESS rounds; however, before doing so we would like to discuss the comparability of the composite scores over countries.

We have seen that the indicators are expressed in the amount of time spent on the different media while the composite scores are just the sum of these amounts

of time. Therefore there is no problem of non-equivalence of the different measures across countries because in all countries these composite scores are defined in exactly the same way in the natural measurement units of time: minutes. This is different from attitudinal concepts, which can have a different interpretation in different countries. So there is no problem of non-equivalence across countries in this case[6].

4 Quality of the composite scores

The way to assess quality of the composite scores (CS_TTM in this case) with formative indicators has been specified by Saris & Gallhofer (2007a) following Bollen (1989) as follows:

$$\text{Quality of CS\_TTM} = 1 - \frac{\text{var}(\delta)}{\text{var}(CS\_TTM)} \qquad (3)$$

Where *"var(CS_TTM)"* is the variance of composite score for TTM; and *var(δ)* is the error variance of the composite score. Equation (3) indicates that the quality of a composite score is equal to 1 minus the proportion of the error variance over the total variance of the composite score. There is no cutoff-point for high and low quality, however the range of this quality index is between 0 and 1, where 1 means perfect measurement without errors and 0 means that there is only error and no substantive variance at all. Equation (3) for the quality requires that we must know the way the composite score is calculated and the error variances and covariances for the different observed variables. The way the composite scores are calculated is indicated in the previous section i.e. as a sum of the observed variables. The error variance *var(δ)* can be computed as follows:

$$var(\delta) = \sum w_i^2\, var(\delta_i) + 2\sum w_i w_j\, cov(\delta_i \delta_j) \quad \text{for i and j = TV, Radio, Newspaper} \qquad (4)$$

where w_i are the weights used to compute the composite score, δ_i denotes the error variance for observed variables which can be obtained separately for each medium. The formula for *var(δ)* simplifies if the error terms are not correlated (no method effects) and becomes the sum of the error variances if the formative model is an unweighted sum. In that case all weights are equal to 1.

The error variances of the observed variables *var(δ)* can be obtained from a Split-Ballot MultiTrait MultiMethod (SB-MTMM) experiments done in the first ESS

[6] We cannot be sure that measures about the definition of politics are comparable across countries. However, with the data we have this possibility cannot be tested.

round. However this experiment is only done for the total time spent on the media. The results for all countries are presented in Appendix 3 (Table A.3).

Using the information in that table and Equation (4) the $var(\delta)$ can be computed. Information from Equation (3) gives the quality of the composite score. An example of this procedure is illustrated for the Austrian case in the website of RECSM[7]. This procedure was followed in order to compute the qualities of the composite score on total time spent in the media (CS_TTM) for the different countries in each ESS round. The results for the comparable countries across time are shown in Table 4.

	CS_TTM			CS_TTPIM			CS_TTEM		
	ESS1	**ESS2**	**ESS3**	**ESS1**	**ESS2**	**ESS3**	**ESS1**	**ESS2**	**ESS3**
Austria	.89	.89	.89	.90	.89	.88	.87	.89	.87
Belgium	.93	.93	.93	.93	.92	.91	.94	.94	.94
Switzerland	.67	.74	.72	.84	.72	.74	.66	.65	.68
Czech Rep.	.78	.77	-	.75	.89	-	.80	.84	-
Germany	.93	.92	.91	.94	.93	.94	.93	.94	.94
Denmark	.80	.80	.79	.85	.86	.84	.80	.82	.82
Spain	.78	.78	.76	.81	.87	.84	.79	.76	.79
Finland	.90	.90	.90	.90	.91	.91	.90	.92	.91
France	.78	.80	.79	.77	.80	.83	.79	.80	.78
Great Britain	.78	.79	.78	.82	.85	.84	.79	.82	.82
Greece	.92	.92	-	.90	.92	-	.94	.95	-
Ireland	.81	.83	.82	.87	.88	.85	.81	.83	.81
Israel	.82	-	-	.84	-	-	.84	-	-
Italy	.83	.82	-	.85	.83	-	.85	.87	-
Netherlands	.86	.86	.85	.89	.89	.90	.86	.87	.87
Norway	.95	.96	.95	.95	.95	.94	.95	.95	.96
Poland	.91	.92	.92	.91	.91	.93	.9	.91	.92
Portugal	.79	.75	.80	.79	.83	.83	.79	.80	.80
Sweden	.71	.68	.72	.73	.75	.73	.66	.69	.70
Slovenia	.76	.74	.77	.82	.81	.80	.75	.78	.76

Table 4: Composite scores quality for ESS1, ESS2 and ESS3

The fact that the SB-MTMM experiments with media variables were only done in the first ESS round makes that for the countries who did not participant in the first round (Cyprus, Bulgaria, Estonia, Slovakia, Turkey, Ukraine, Romania, Russia, Iceland, Luxemburg, Hungary[8] and Latvia) a value for composite score

[7] "Quality" in *http://www.upf.edu/survey/working/extrafiles*

[8] Hungary participated in the ESS1 but participated not in the SB-MTMM experiment

quality cannot be computed. The reason is that, as aforementioned, additional information about SB-MTMM is necessary for computing these qualities.

Missing values (-) in Table 4 are due to the lack of participation of such country in the round; for instance, Israel participated in the first round only, Italy and Czech Republic in the first and second rounds.

For the composite scores CS_TTPIM and the CS_TTEM we use the same information with respect to the quality of the different indicators obtained in the analysis of the SB-MTMM experiment. We consider that the proportion of the error that one can make in the different media questions with the same scale is the same. Since the quality of each indicator is equal to:

$$\text{Quality} = \left[1 - \left(\text{var}(\delta_i) / \text{var}(CS\_TTM)\right)\right] \quad (5)$$

the error variance *var(δ)* can be estimated. Details for this procedure are explained on the web of RECSM[9]. Next, the quality of the composite score can be estimated in the same way using Equation (5). The qualities for CS_TTPIM and CS_TTEM are shown in the last columns of Table 4.

The results show mainly that the quality of the composite scores is roughly stable across time for the different countries. The differences are minimal for each composite score. We can realize that, in general, the quality for these composite scores is quite good and the quality for "Total Time spent in Political Issues in the Media" (CS_TTPIM) is slightly better than the other two. Larger is the difference across countries; for instance in ESS round 1 for the CS_TTM, there exists a difference of .28 between Switzerland (quality equals to .67) and Norway (quality equals to .95).

This table shows that the quality of the composite scores in all countries is quite good but there is considerable variation in quality, which means that correction for measurement error in the analysis is necessary in order to compare relationships with other variables across countries. How this can be done will be shown in the next section.

[9] "Error Variance" in *http://www.upf.edu/survey/working/extrafiles*

5 Validity of the concept by postulation "Total Time"

5.1. Literature review

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

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.

There are several other studies relating media and political participation. We can mention studies by Aarts & Semetko 2003; Eveland & Scheufele 2000; Newton 1999; Swyngedown & Goeminne 2007; Vreese & Semetko 2004. A relationship between media use and knowledge has also been reported (Guo 2000; Vreese & Boomgaarden 2006).

This brief overview of the literature suggests that many studies detected a relationship between media use and political discussion, political participation and political knowledge. But most of the time a relationship is indicated for a specific type of media and only occasionally the total time spent on the media is used as a variable. However, in this section we would like to test whether the total time spent on the traditional media has an effect on those political variables or if one indeed should differentiate for the different media. If the latter would be the case, the total time concepts are irrelevant and one can forget about these concepts and just use the measures for time spent on the different media.

Figure 5 specifies the idea that "total time" influences the political variables "political discussion", "political knowledge" and "political participation". For the measurement of these political variables we refer to Appendix 4. Figure 5 shows that we start with the specification of a model where the political variables are only directly influenced by the total time variable and not by the specific media use variables. However, we do not expect the total time variable to be the only

explanatory variable. Therefore, we assume correlated disturbance terms between the dependent variables. The MIMIC model in Figure 5 will be used for multiple groups using structural equation (SEM) model software for 20 countries from the ESS first round as an example for external validity of the model on media use. In Figure 5 the errors for the endogenous observed variables are indicated as ε_i, while disturbance terms for the endogenous latent variables are denoted as ζ_η.

For the evaluation of the structural model we will use the procedure of Saris et al. (2009). They argue that the standard testing procedures and fit measures do not provide proper information about the quality of models. They have suggested a procedure, which looks directly for misspecifications in the model. To test for misspecifications the Expected Parameter Change (EPC), Modification Index (MI) and the power of the test have to be known. EPC gives direct estimates of the size of the misspecification for all fixed parameters, while the MI provides the test statistic for the estimated misspecification (Saris et al. 1987). However, these two statistics are not sufficient for determining misspecifications because the MI depends on other characteristics of the model. For this reason, the power of the test is also necessary. So, testing procedure of Saris et al. specifies on the basis of the size of MI, the power of the test and the EPC whether a restriction is most likely incorrectly made (i.e. a misspecification).

In this specific case we are particularly interested in the direct effects of the time spent on the different media on the political variables.

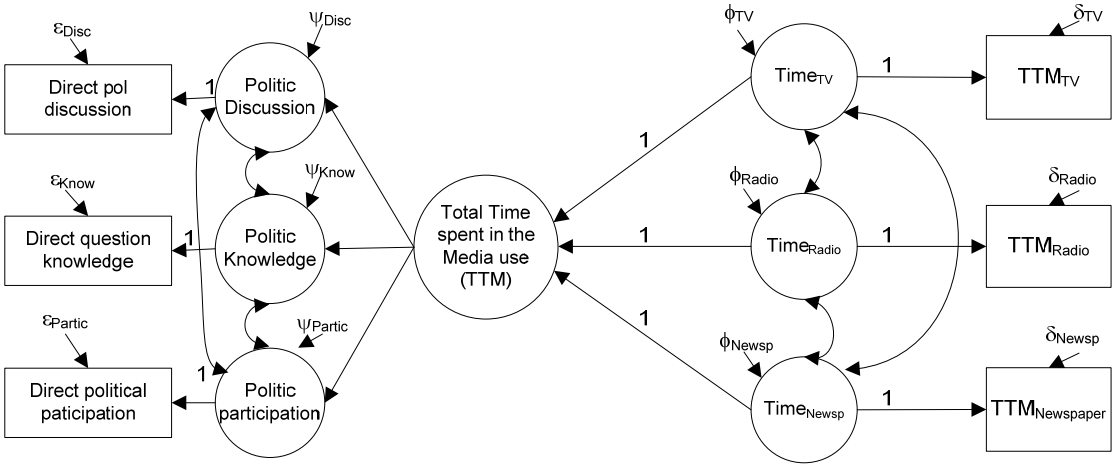

Figure 5: TTM model with additional observed variables

The results of the estimation and testing on misspecifications of this model will be discussed in sequences for the different purposes of media use. We will start with media use for political issues.

5.2. Results for Total Time spent on Political Issues in the Media use (TTPIM)

Using the testing procedure of Saris et al. (2009) the TTPIM model requires additional effects in most of the countries from "total time spent on political issues in the newspaper" on "political discussion" and "political participation". These effects are presented in Figure 6 as arrows with a positive sign, which means that the additional effect is positive. After introducing these direct effects no more misspecifications are suggested in the model. Thus, the final TTPIM model is shown in Figure 6.

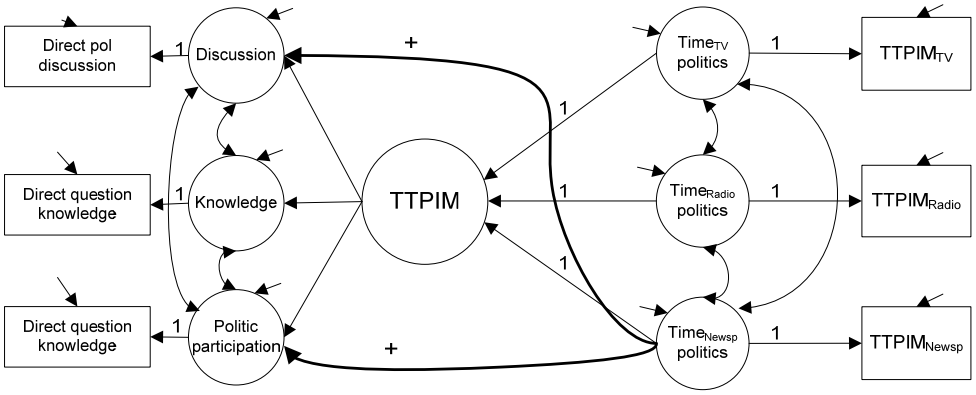

Figure 6: TTPIM model with additional observed variables

The effects from the TTPIM latent variable on the political variables are shown in the first three columns in Table 5. Additional effects from "reading political issues in the newspaper" on "political discussion" and "participation" have been found, which are shown in the last columns of Table 5. In 14 countries at least one of these effects is significant; in 6 countries none of these effects is significant. On the other hand in 19 out of 20 countries at least one of the direct effects of newspaper reading is significant.

Table 5 shows that the additional effect "newspaper" on "political discussion" is significant in all countries, except Ireland; and "newspaper" on "political participation" is significant for 15 countries. In the Irish case, the effects of the latent variable TTPIM on the political variable are the only significant ones; there is no direct effect from reading newspapers. In Austria, Finland, Portugal, France, Spain and Sweden the use of the latent variable TTPIM does not make sense because all effects on political participation and discussion come from reading political issues in the newspapers and none of the effects of TTPIM is significantly different from zero.

It is important to note that the direct effects from the latent variable TTPIM on the political variables are positive (except for Norway on participation[10]). This means that the more time they spent on the media for political issues, the more they tend to participate, to know political issues and to discuss with other people about political topics.

	TTPIM on knowledge	TTPIM on discussion	TTPIM on participation	Newspaper on discussion	Newspaper on participation	Sample size
Austria	.094	.109	-.078	**.341***	**.218***	2257
Belgium	.009	**.142***	**.090***	**.357***	**.133***	1898
Switzerland	**.185***	**.199***	.045	**.231***	**.137***	2040
Czech Republic	.005	**.324***	.052	**.163***	.055	1360
Germany	**.068***	**.196***	**.028***	**.237***	**.118***	2833
Denmark	**.121***	-.058	-.043	**.275***	**.118***	1506
Spain	.024	-.009	-.081	**.173***	.092	1729
Finland	.000	.000	.000	**185***	**. 130***	2000
France	.095	.002	-.005	**.524***	**.211***	1360
Great Britain	**.109***	.101	**.130***	**.403***	**.120***	2052
Greece	.006	**.378***	.040	**.211***	.088	2566
Ireland	**.165***	**.407***	.060	-.046	.037	1922
Israel	.015	**.243***	-.066	**.244***	**.054***	2499
Italy	**.115***	.148	.056	**.465***	**.206***	1207
The Netherlands	**.078***	.096	**.075***	**.214***	**.095***	2364
Norway	.000	.000	**-.118***	**.289***	**.237***	2036
Poland	**.073***	**.261***	.008	**.462***	**.219***	2110
Portugal	.000	.000	.000	**.262***	**.138***	1511
Sweden	.000	.000	.000	**.417***	**.145***	1999
Slovenia	**.075***	.025	.004	**.454***	.081	1519

Table 5: Standardized estimates for the effects of TTPIM plus the extra effects needed to correct for misspecifications in the model. *Significant effects.

These results suggest that newspaper reading of political issues has a direct effect on political discussions and political participation independent of the other two media. Therefore two extra direct effects had to be introduced.[11]

[10] In some countries the effects were negative but not significantly different from zero. In such cases we fixed the coefficients to zero. This was not the case in Norway. In Norway, the correlations between the observed variable political participation with television and radio are negative, while with newspaper were positive. This is the reason for the negative effect.

[11] We also tried to test for an interaction effect of newspapers reading and total time on political variables but this model did not show that the interaction effect was needed.

5.3. Results for Total Time spent on Entertainment in the Media use (TTEM)

The next hypothesis we want to test is whether the use of the latent variable TTEM makes sense. In that case this variable should have significant negative effects on the same political variables and the model should not be misspecified by omission of direct effects of the different media use on the political variables.

In order to obtain the $TTEM_i$ measures, we subtracted the time spent on political issues ($TTPIM_i$) from the total amount of time spent in the media (TTM_i).

The result is the total time spent in the media for different purposes than on political issues; therefore "total time spent on entertainment" (TTEM) is defined as an aggregation of the topics in Figures 1 and 2 (entertainment, business, sport, hobbies, education and the arts).

The model for these variables was specified in the same way as for the variable TTPIM. Misspecifications in the model were determined as before.

The results for the model suggested additional effects in the TTEM model. These additional effects are from "television watching" on "political discussion" and "political participation". These effects are presented in Figure 7 as arrows with a negative sign, which means that the additional effect is negative. These lines are dashed because the effect is only needed for 6 countries and not for all countries. The final TTEM model is shown in Figure 7 and standardized estimates in Table 6.

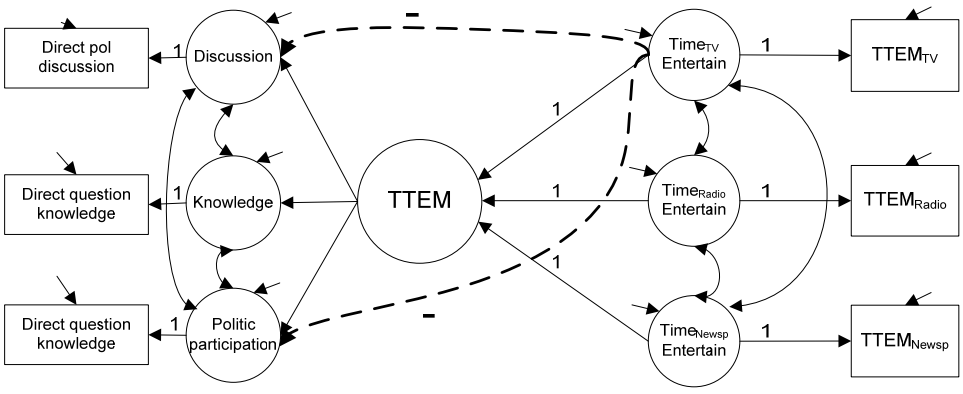

Figure 7. TTEM model with additional observed variables

The effects from the latent variable TTEM on the political variables are shown in the three first columns in Table 6. The additional effects from "television" on "political discussion" and "participation", which have been found looking for misspecifications in the model, are shown in the two last columns of Table 6.

	TTEM on knowledge	TTEM on discussion	TTEM on participation	television on discussion	television on participation
Austria (AT)	**-.140***	**-.334***	**-.231***	0	0
Belgium (BE)	-.069	**-.143***	**-.201***	**.268***	0
Switzerland (CH)	-.072	.142	-.001	**-.653***	**-.395***
Czech Republic (CZ)	-.052	**-.247***	**-.148***	0	0
Germany (DE)	**-.067***	.041	-.014	**-.382***	**-.276***
Denmark (DK)	-.057	**-.294***	**-.195***	0	0
Spain (ES)	-.145	**-.216***	-.042	0	0
Finland (FI)	**-.081***	**-.105***	-.048	**-.172***	**-.155***
France (FR)	.047	**-.349***	**-.223***	0	0
Great Britain (GB)	**-.104***	**-.477***	**-.226***	0	0
Greece (GR)	-.968	**-.342***	-.056	0	0
Ireland (IE)	**-.155***	**-.219***	**-.156***	**-.264***	0
Israel (IL)	-.043	**-.355***	**-.160***	0	0
Italy (IT)	-.057	**-.288***	**-.133***	0	0
The Netherlands (NL)	**-.109***	**-.292***	**-.162***	**-.269***	**-.134***
Norway (NO)	-.035	**-.311***	**-.191***	0	0
Poland (PL)	-.035	**-.362***	**-.162***	0	0
Portugal (PT)	-.076	**-.111***	.009	0	0
Sweden (SE)	**-.162***	**-.502***	**-.201***	0	0
Slovenia (SI)	**-.111***	-.097	**-.115***	0	0

Table 6: Standardized values for the effects of TTEM plus the extra effects needed to correct for misspecifications in the model. *Significant estimates

The results in Table 6 show that direct effects from TTEM latent variable on political variables are all negative. In 19 countries there are significant effects from TTEM on the political variables. For 14 countries, there are only negative direct effects, which mean that this model indeed indicates the influence of entertainment on political issues. The effect of TTEM has a larger negative influence on political discussion than participation or knowledge.

The additional effects were necessary in six countries. These were Belgium, Finland, Germany, Ireland, the Netherlands and Switzerland. The largest negative effects are from television to political discussion.

Switzerland is the only country where the negative effects on the political variables come only from total time watching television about entertainment (non-political) issues. Therefore, no direct effect from the TTEM is found.

In sum, the more time people spend on the media for entertainment (non-political issues) the less they participate in politics, discuss about political issues and the less knowledge they have about political topics. The negative effect on political discussion is again the largest.

5.4. Results for Total Time spent on the Media (TTM)

In the case of the TTM model we again tested whether a single latent variable composed by the total amount of time spent on the media (time spent on political issues and time spent on non-political issues) has influence on the political variables discussion, knowledge and participation.

The model for the total amount of time spent on the media (TTM) is specified also in Figure 5. Additional effects are checked using the Saris et al. (2009) procedure. The final model for TTM is shown in Figure 8 where additional effects are needed from "television" and "newspaper" on "political discussion" and "political participation". Direct effects from newspaper on political variables are positive and from television are negative. This is shown in Figure 8 where the positive effects are arrows with a positive sign while negative effects are arrows with a negative sign.

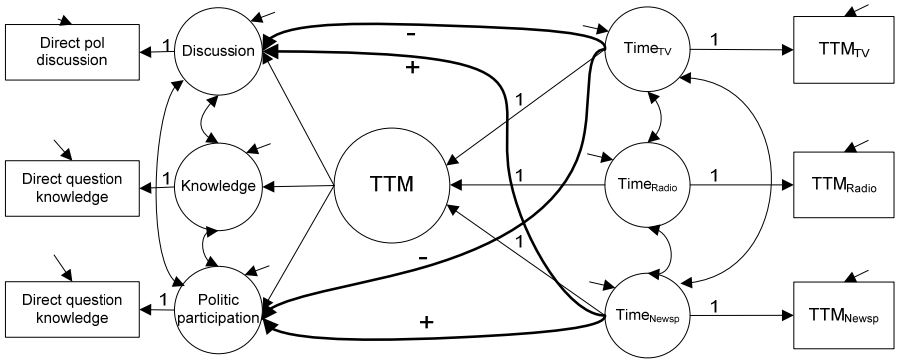

Figure 8: TTM model with additional observed variables

The effects from the latent variable TTM on the political variables are shown in the three first columns in Table 7. Additional effects from "television" on "political discussion" and "political participation" are in the next two columns, and the last three columns of Table 7 are effects from reading newspapers. Table 7 shows that there is no clear effect from the latent variable TTM on the political variables we are interested in. Once additional effects have been specified, there are almost no significant effects from TTM on political variables.

Additional paths needed are negative direct effects from television on "political discussion" and "political participation", positive effects from newspapers on "political discussion" and "political participation" and also on "political knowledge" for ten countries.

The direct effects from television and newspaper means that each medium has a different behaviour, there seems not to be a need for a common variable "total time spent in the media" because this variable has hardly any significant effects on the political variables. This can be explained in the measure because different purposes and interests are mixed together.

	TTM on knowledge	TTM on discussion	TTM on participation	TV on discussion	TV on participation	NP on discussion	NP on participation	NP on knowledge
Austria	**-.172***	-.039	**-.146***	**-.115***	**.091***	**.315***	**.149***	**.354***
Belgium	.021	-.008	-.003	**-.198***	**-.137***	.071	.006	0
Switzerland	-.076	.030	-.081	**-.300***	**-.214***	**.259***	**.124***	**.244***
Czech Rep.	.019	**.247***	.012	**-.313***	**-.245***	**.213***	.062	0
Germany	-.053	**.169***	.046	**-.377***	**.343***	**.285***	**.087***	0
Denmark	.090	**-.113***	.024	**-.194***	**-.237***	**.253***	**.104***	0
Spain	**.183***	.090	**.108***	**-.153***	**-.178***	**.286***	**.112***	**.261***
Finland	-.091	-.051	-.054	**-.238***	**-.195***	**.227***	**.187***	0
France	.110	**.214***	.079	**-.318***	**-.226***	**.122***	.059	0
Great Br.	-.112	.056	.049	**-.347***	**-.248***	**.207***	**.101***	**.246***
Greece	**.153**	-.068	.094	.077	**.157***	**.526***	**-.164***	**.277***
Ireland	**-.154***	-.045	**-.096***	**-.218***	**-.090***	**.313***	**.137***	**.340***
Israel	.059	.084	.039	-.034	**-.172***	**.271***	.018	0
Italy	.008	-.050	-.035	.053	.003	**.732***	**.325***	**.566***
Netherlands	-.077	-.029	-.005	**-.373***	**-.314***	**.258***	**.158***	0
Norway	.089	-.116	**-.118***	**-.239***	**-.164***	**.233***	**.138***	0
Poland	**.153***	.050	.025	-.003	**.096***	**.276***	**.170***	**.198***
Portugal	**.240***	**.153***	.056	-.101	**-.146***	**.315***	**.266***	**.341***
Sweden	**.176***	**.119***	.087	**-.286***	**-.172***	**.257***	**.120***	**.264***
Slovenia	.078	-.021	-.097	-.002	-.010	**.372***	**.141***	0

Table 7: Standardized values for the effects of TTM plus the extra effects needed to correct for misspecifications in the model. *Significant estimates

This model even requires additional positive effects from newspaper on political knowledge for ten countries, which are not needed in the TTEM or TTPIM models. This means that once we aggregate these two models, we are not creating an adequate construct to study its relationship with political variables and probably also other variables.

6 Conclusions

This paper first presents the construction of the media module of the European Social Survey (ESS). With this information we could define the concepts by postulation ”Total time spent on political issues presented in the media”, the “Total time spent on entertainment (non-political issues)” and the aggregation of these two measures which is the “Total time spent on the media for all purposes”. The composite scores for these variables are defined as the sum of the time spent for the different purposes in the three traditional media; TV, Radio and newspapers. Since the scales of these behavioural variables are the same, there is, in this case, no problem of non-equivalence of the measures.

However one can wonder whether the concepts are needed. It is possible that effects on other variables do not come from the total time spend on all media but from the time spent on different media. This would be the case if the effect of the different media is different— then the total time variable across the different media would not be necessary.

In order to test this idea the same structural model was specified for each of these three variables (TTPIM, TTEM and TTM) adding three political variables (knowledge, discussion and participation) as dependent variables. Data is from the first round of the ESS. We expected a positive effect of TTPIM on political variables and a negative one from the TTEM construct. The effects from TTM were not clear due to the aggregation of these two variables, which indicate different purposes of media use.

The results showed that TTPIM and TTEM, despite some additional effects, can be used because these variables have significant effects on the political variables. We found a positive influence from TTPIM (including positive direct effects from newspapers on political participation and discussion) and negative effect from TTEM (including negative direct effects from television on political participation and discussion).

The result for the "Total time spent in the media"-TTM- (which is the sum of the other two) shows that it leads to problems if one tries to relate the exposure to the media with political variables and possibly also with other variables. The reason for these problems is that the purposes of media exposure can be very different and, therefore, the aggregated time over the different exposures contains many different components with different relationships, which will confuse the analysis. In order to get a fitting model with total exposure required, many direct positive effects from newspaper reading and negative ones from television watching were necessary in the model while the effects from the construct on those political variables were mostly non significant. This shows that in this case the use of time spent on different media should be utilized and not the aggregated score over all three media. It should be mentioned that this is a different issue than mentioned by Potter (1997), who suggested that the measures for specific purposes are more precise than measures of the overall exposure to the media.

Given these results we have brought the scores for all respondents on the variables TTPIM and TTEM in a file, which can be obtained from the authors for further research.

Finally, we would like to include a disclaimer. So far, only the traditional media are taken into account. For future research, the ESS should consider the inclusion of similar measures for the new media. Presently, these concepts by postulation can only represent measures of the use of traditional media. If one would like to measure the effects of general media use, one should also include the new media in the data collection. This should be done in the same way as it is done for the traditional media.

Part 2 The Measurement of Interest in Political issues presented in the Media (IPIM)[iii]

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).

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 do not watch 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 programmes 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

Figure 1: Questions for the media use module

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:

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

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 2007a) 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.

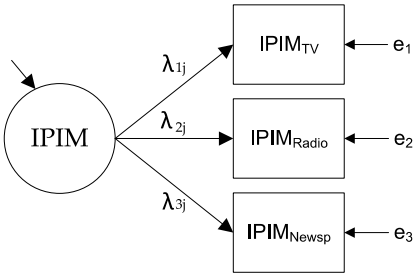

Figure 2: The measurement model for IPIM

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.

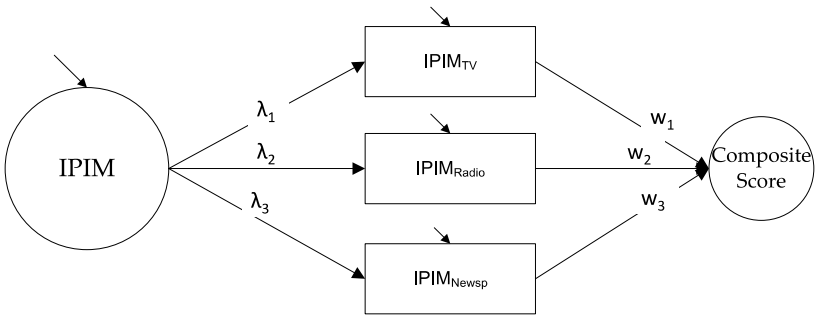

Figure 3: The construction of an index for IPIM as a composite score

In Figure 2 and 3 the λ_i is the factor loading for each reflective indicator, which can also be interpreted as the quality coefficient of specific indicator, w_i is the

weight used for computing the composite score and e_i is the measurement error in indicator i [12].

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 1st, 2nd, 3rd and 4th ESS rounds.

Finally, substantive relationships will be studied with some outcome variables, which are related to IPIM and political interest according to our hypotheses, derived from the literature.

7 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 respondents were asked directly for a numeric estimate of the amount of time spent on each medium[13]. 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 Research and Expertise Centre for Survey Methodology (RECSM)[14]. The results were, for instance, that the numerical values obtained in Austria for the 8th category are 234 minutes for television, 271 for radio and 180 for newspaper. For Spain they are 261, 300 and 257 or for Belgium they are 266, 300 and 300, respectively.

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

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

[14] See part 1 and *http://www.upf.edu/survey/working/extrafiles*

The newly created numeric values of these variables can be found on the website of RECSM[15]. Table 1 presents the transformation to minutes.

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

Table 1: Original and transformed media use measures

The values obtained in each country are also used in the next ESS rounds. However, the participating countries in the next 3 ESS rounds were not exactly the same as the ones in the 1st round. For instance, Estonia started to participate from the 2nd 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, for each new country, we specified for the 8th category the amount of minutes of a country from the 1st round, which has the most similar distribution for each medium (Appendix 2.2).

8 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 (2007b), & Saris (2007) and Oberski, Saris & Hagenaars (2010) 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.

8.1 Invariance of the measurement instruments

The commonly[16] used criteria for measurement invariance are configural, metric and scalar invariance (Horn et al. 1983; Meredith 1993; Steenkamp &

[15] "Scores" in *http://www.upf.edu/survey/working/extrafiles*

Baumgartner 1998). In order to better understand these criteria, we generalize the reflective model of Figure 2 by creating Figure 4 which are also represented in the equations and assumptions presented at the left hand side of the figure.

$y_1 = \tau_1 + \lambda_{1j}\eta_j + e_1$

$y_2 = \tau_2 + \lambda_{2j}\eta_j + e_2$

$y_3 = \tau_3 + \lambda_{3j}\eta_j + e_3$

$E(e_i) = E(e_i\ \eta_j) = E(e_i\ e_j) = 0$ for $i \neq j$

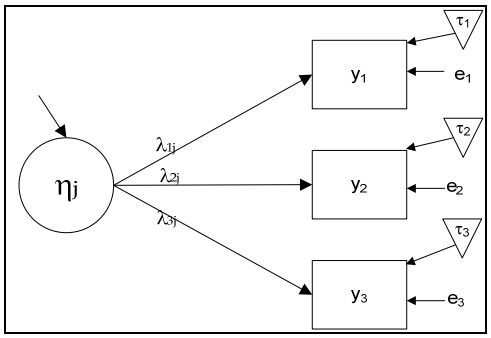

Figure 4: reflective model with intercepts (τ) and loadings (λ)

In this model "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.

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 modelling (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.

[16] 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 2007a) but in this case we can not make this distinction.

8.2 Testing for measurement invariance in ESS

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 without 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 (λ_{IPIMtv} in our case) and the remaining loadings are estimated but each one must be equal across countries. We include to multiple group analysis all countries from four ESS rounds, which makes together 97 groups. It occurs that metric invariance does hold for all 30 participating countries in four rounds. This was detected using programme JRule, which has been developed by Saris, Satorra & Van der Veld (2009). This program determines whether misspecifications are present in the specified model taking into account the power of the test, which in these cases was very high (>.9). For more details, including JRule starting specifications, we refer to the website of the RECSM[17]. This result means that these indicators are metric invariant, which makes comparison of relationships across countries possible[18].

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 RECSM[19]. It turned out that across countries only for Slovenia there exists some significantly differing intercepts for IPIM$_{TV}$. An extra test confirmed that for Slovenia the IPIM$_{TV}$ was scalar invariant across time. Therefore the intercepts were constraint to be equal in all rounds for Slovenia but possibly different from those of the other countries.

The final results of the two tests are shown in Table 2. In the first row of Table 2 the estimates for the 29 scalar invariant countries are mentioned. The second row presents the estimates for the Slovenia.

[17] "Detection of deviating loadings" in *http://www.upf.edu/survey/working/extrafiles*

[18] A former study by Coromina, L., Saris, W.E., and Oberski, D. (2008) lead to partal invariance, but that was due to a different misspecification detection approach.

[19] "Detecting deviating intercepts" in *http://www.upf.edu/survey/working/extrafiles*

	$\lambda_{tv,IPIM}$	$\lambda_{rd,IPIM}$	$\lambda_{nwsp,IPIM}$	τ_{tv}	τ_{rd}	τ_{nwsp}
29 Invariant countries	1	.801	.56	0	.066	.388
Slovenia	1	.801	.56	0	.066	*.295*

Table 2: Scalar invariance and final estimates for the differing countries in round 1-4

Figure 5 shows graphically the deviation from the scalar invariant relationships presented in Table 2, the blue line stands for $IPIM_{tv}$, the red one for $IPIM_{rd}$, the green one for $IPIM_{nwsp}$. The dashed green line is for Slovenian $IPIM_{nwsp}$, which as mentioned, is the only differing country, and it only deviates by intercept for $IPIM_{nwsp}$, which is about .1 lower than in the other countries.

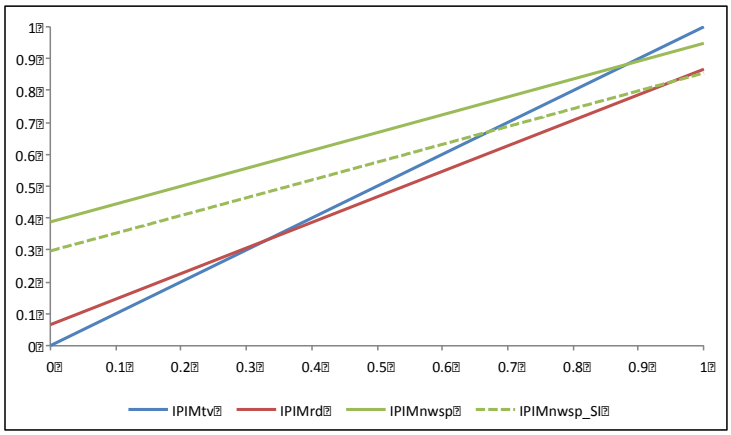

Figure 5: Graphical scalar invariance and deviations

It has been shown that there are some statistically significantly different coefficients in Slovenia, 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 (2007a) 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 observed deviation creates serious problems in the comparison across countries.

4. 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.

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 (2007a). As this is true for all countries except Slovenia, we expect on statistical grounds for all other countries, that the means of the composite score will be statistically comparable with the latent variable means.

The simplest way is to use for the composite score the unweighed mean, but on the other hand the weighted mean could lead to better quality (Saris and Gallhofer 2007a). Next we compare both possibilities, starting from unweighed approach.

4.1 The unweighed composite score

We have made the comparison between the latent means and the means obtained using the unweighed 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 3 presents the two sets of means for all countries in cumulative data of first four ESS rounds.

Country	Latent means	Unweighted Composite scores means
Austria	0.428	0.495
Belgium	0.396	0.460
Bulgaria	0.352	0.431
Switzerland	0.445	0.502
Cyprus	0.271	0.371
Czech Republic	0.340	0.428
Germany	0.396	0.476
Denmark	0.516	0.555
Estonia	0.411	0.476
Spain	0.416	0.490
Finland	0.484	0.530
France	0.433	0.496
United Kingdom	0.361	0.424
Greece	0.286	0.387
Hungary	0.357	0.432

Country	Latent means	Unweighted Composite scores means
Ireland	0.412	0.464
Iceland	0.492	0.520
Israel	0.482	0.538
Italy	0.407	0.465
Luxembourg	0.436	0.505
Netherlands	0.452	0.505
Norway	0.555	0.582
Poland	0.407	0.461
Portugal	0.480	0.533
Russian Federation	0.330	0.411
Sweden	0.497	0.533
Slovenia	0.439	0.460*
Slovakia	0.358	0.428
Turkey	0.378	0.449
Ukraine	0.390	0.461

Table 3: **Comparison of means for cumulative date from 4 ESS round (***not scalar invariant countries)

This table shows that the means of the composite scores are in all countries higher than the latent means. This follows from the fact that the indicators do not have all a slope of 1 and an intercept of zero (see table 2). As a consequence it can be shown that the relationship between the two means is a linear with the following values :

Mean (CS) = .151 + .72 mean(latent)

Because this relation is the same for all countries the means across countries can be compared any way.

The correlation between the latent means and the mean of the composite score without Slovenia is, in the cumulative data of ESS, .989. And the correlation with all 30 countries, including the Slovenia is .981. 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 plots for ESS 1 and ESS 4 (Figures 6 and 7).

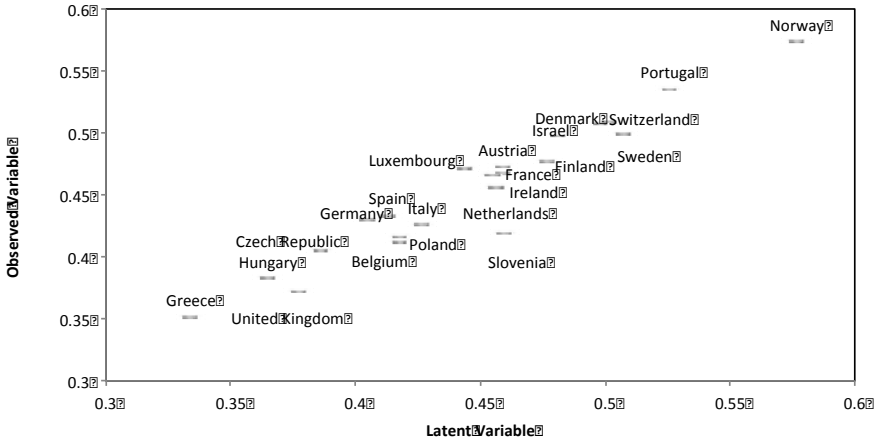

Figure 6: The relationship between the means of the Latent variables and composite scores for the 1st ESS round

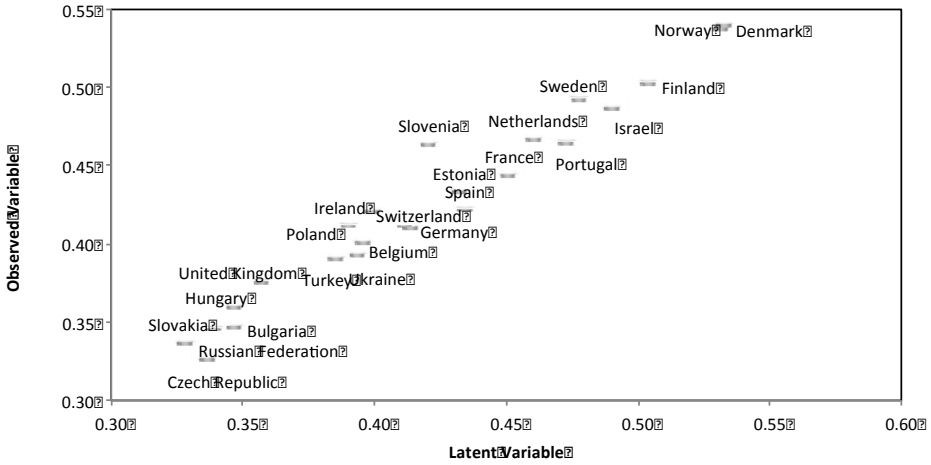

Figure 7: The relationship between the means of the Latent variables and composite scores for the 4th ESS round

Figures 6 and 7 shows that the non scalar invariant Slovenia is deviating a bit from the others, but it can not be considered as a serious outlier.

4.2 The weighted composite score

As mentioned, when metric invariance requirement across time and between countries is fulfilled, there is also a possibility to use weighed composite scores. The literature refers to three different kinds of weights: regression weights, Barlett weights and Anderson and Rubin weights (Saris and Gallhofer 2007a). In this particular case all these three approaches lead to an identical outcome, so we will use the first one (regression weight).

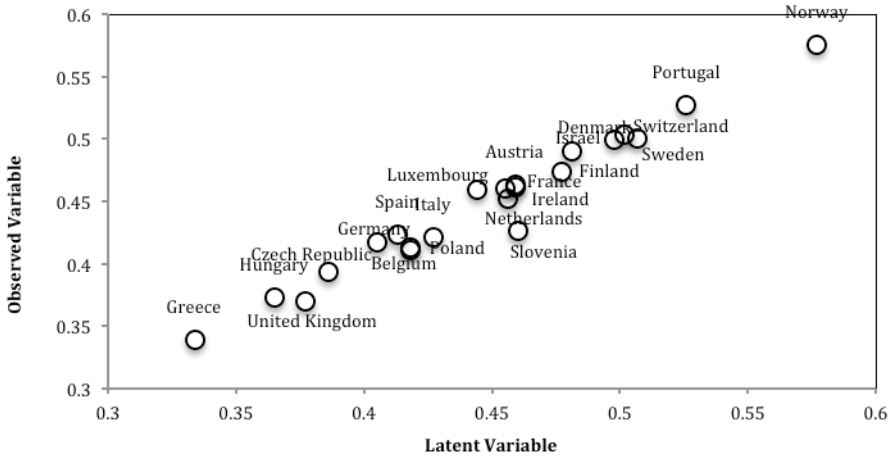

Figure 8: The relationship between the means of the Latent variables and weighed composite scores for the 1st ESS round

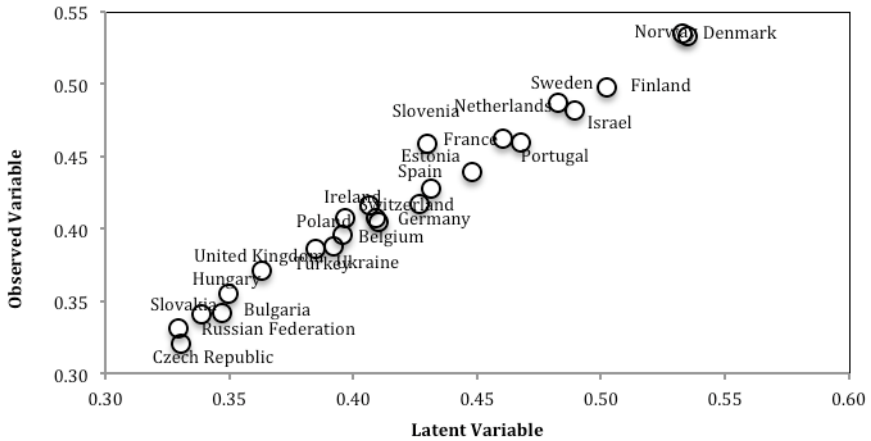

Figure 9: The relationship between the means of the Latent variables and composite scores for the 4th ESS round

As expected, the correlations for weighted composite scores are even higher than for un-weighted scores. In the cumulative data of ESS the correlation without Slovenia is .997. And the correlation with all 30 countries, including the Slovenia is

.992. This means, that the bias in the means is again very small. The scatter plots in Figure 8 and 9 confirm this also.

We can see, that the means of weighted and unweighted composite scores, both for first (Figures 6 and 8) and for forth ESS round (Figures 7 and 9), are very similar, which means that from this point of view weighing doesn't have much influence on measures.

4.3. Means of composite scores

The means can be directly interpreted in the proportion of the time spend on political issues in the media. So we see that in year 2002 in Norway, Denmark, Switzerland, Portugal and Sweden more than 50% of the time spent on the media is dedicated to political issues (Figures 6 and 8). Therefore, the people in these countries were 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 were the least interested in political issues in the media; they use the media less than 45% of the time for information on political issues.

When comparing means with year 2008 (Figures 7 and 9), we can see several changes, but the larger structure is the same – most interested in political issues in media are Scandinavians and less interested are Czechs, Slovaks, Russians and Bulgarians, clearly under 40% line are also Brits and Hungarians.

So we can conclude that political interest in traditional media is rather stable across time, which means that media consumption habits don't change fast.

4.4. 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 2007a). We consider this restriction too strict; we demonstrate in the previous section that the country which is not invariant, for different rounds, does not have a large deviation in the composite score means. This is so because the differences across countries were 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 example, in all rounds for Slovenia, newspaper was the differing medium. Based on the results presented in table 2 the bias in the mean is calculated to be .03 because:

$(\tau_{\text{scalar invariance}} - \tau_{\text{slovenia}})\ w_{nwp} = [(.388 - .295)\ (1/3) = 0.03$

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.

4.2. 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. 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 quality of the composite scores is shown in Figure 10.

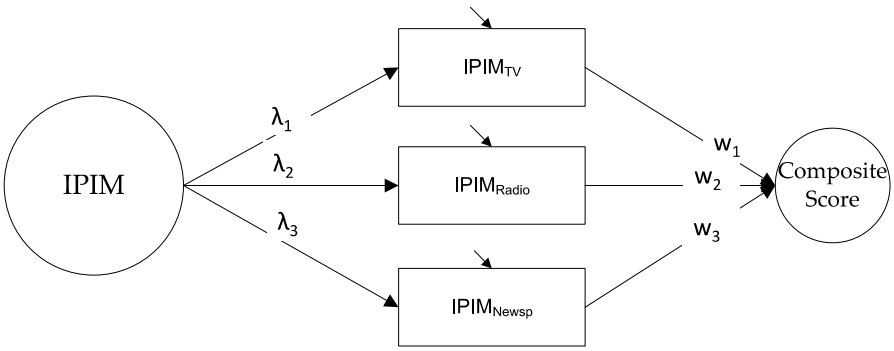

Figure 10: Reflective model and composite score for IPIM

So 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 2007a):

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

Where "λ_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 and "var(S)" is the

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

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

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

To give an example, for unweighed composite score of Austria in the 1st round $r(IPIM, S)_{AU}$ = .59 because λ_1= .54; λ_2= .36 ; λ_3= .22 are the standardized loadings in Austria for this model. Then, the quality is $r^2(IPIM, S)_{AU}$ = .34, Using the same procedure for the remaining countries, we obtain the qualities for the composite scores for all countries. The qualities for the weighed and unweighted composite scores in all four round and cumulatively are shown in Table 4.

Quality	ESS1		ESS2		ESS3		ESS4		Mean1		Mean2	
	Q_1	Q_2	Q_1	Q_2	Q_1	Q_2	Q_1	Q_2	Q_1	SD	Q_2	SD
Austria	0.34	0.36	0.35	0.37	0.32	0.33	-	-	0.34	0.02	0.35	0.02
Belgium	0.33	0.36	0.36	0.37	0.32	0.34	0.32	0.33	0.33	0.02	0.35	0.02
Bulgaria	-	-	-	-	0.33	0.34	0.36	0.39	0.34	0.02	0.36	0.03
Switzerland	0.32	0.33	0.34	0.35	0.31	0.32	0.35	0.36	0.33	0.02	0.34	0.02
Cyprus	-	-	-	-	0.37	0.39	0.44	0.48	0.40	0.05	0.44	0.06
Czech Republic	0.37	0.40	0.38	0.41	-	-	0.35	0.38	0.37	0.01	0.39	0.01
Germany	0.39	0.41	0.40	0.41	0.33	0.34	0.35	0.37	0.37	0.03	0.38	0.03
Denmark	0.37	0.39	0.36	0.37	0.32	0.33	0.35	0.37	0.35	0.02	0.36	0.02
Estonia	-	-	0.43	0.45	0.32	0.33	0.35	0.38	0.37	0.06	0.39	0.06
Spain	0.32	0.34	0.32	0.33	0.30	0.31	0.32	0.34	0.31	0.01	0.33	0.01
Finland	0.38	0.40	0.40	0.41	0.34	0.35	0.35	0.36	0.37	0.03	0.38	0.03
France	0.31	0.33	0.34	0.35	0.29	0.30	0.31	0.33	0.31	0.02	0.33	0.02
United Kingdom	0.33	0.35	0.36	0.38	0.30	0.32	0.31	0.33	0.32	0.03	0.34	0.03
Greece	0.37	0.40	0.42	0.44	-	-	0.39	0.42	0.39	0.02	0.42	0.02
Hungary	0.37	0.40	0.38	0.40	0.32	0.34	0.31	0.33	0.34	0.03	0.37	0.04
Ireland	0.24	0.26	0.26	0.27	0.27	0.28	0.28	0.29	0.26	0.01	0.27	0.01
Iceland	-	-	0.34	0.35	-	-	-	-	0.34	-	0.35	-
Israel	0.28	0.30	-	-	-	-	0.24	0.25	0.26	0.03	0.27	0.03
Italy	0.31	0.34	0.36	0.38	-	-	-	-	0.33	0.03	0.36	0.03
Luxembourg	0.30	0.31	0.36	0.38	-	-	-	-	0.33	0.05	0.34	0.05
Netherlands	0.35	0.37	0.35	0.37	0.31	0.32	0.31	0.32	0.33	0.03	0.34	0.03
Norway	0.37	0.39	0.37	0.38	0.30	0.32	0.33	0.34	0.34	0.03	0.36	0.03
Poland	0.36	0.39	0.35	0.37	0.30	0.31	0.33	0.35	0.33	0.03	0.35	0.03
Portugal	0.33	0.35	0.33	0.34	0.27	0.28	0.33	0.34	0.31	0.03	0.33	0.03
Russian Federation	-	-	-	-	0.34	0.36	0.36	0.38	0.35	0.01	0.37	0.02
Sweden	0.34	0.36	0.32	0.33	0.29	0.30	0.28	0.29	0.31	0.03	0.32	0.03
Slovenia	0.31	0.33	0.31	0.32	0.26	0.27	0.27	0.29	0.29	0.03	0.30	0.03
Slovakia	-	-	0.36	0.38	0.32	0.33	0.35	0.38	0.34	0.02	0.36	0.03
Turkey	-	-	0.37	0.39	-	-	0.36	0.37	0.36	0.01	0.38	0.01
Ukraine	-	-	0.35	0.36	0.32	0.33	0.30	0.32	0.32	0.02	0.34	0.02
Total	0.32	0.34	0.34	0.35	0.30	0.31	0.31	0.32	0.32	0.31	0.33	0.02

Table 4: Quality for Composite Scores

Q_1 stands for the quality of the unweighted and Q_2 for quality of the regression weighted composite scores. 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 table shows that the difference in quality between the un-weighted and weighted composite scores for IPIM is very small (1%). Given this situation one can make the choice for the simpler un-weighted composite score.

The qualities in all countries are rather low. The reason is that the correlations between the observed variables for the different media are also not very high.

We can also see that there is, as can be expected, a difference in the quality estimates across the different rounds. This variation is summarized in the standard deviation (SD). The standard deviation varies between .01 and .06. The large values have been obtained for Estonia, Cyprus and Luxembourg, which have also not participated in all rounds. The mean quality of the un-weighted composite score is .32 and in most countries the deviation from this value is within the confidence interval of countries mean. Therefore we suggest that we fix the quality for all countries on .32 i.e. the explained variance in the calculated score for IPIM is only for 32% explained by the variable to be measured "Interest in political issues in the media".

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 scores 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 2007a) and will be demonstrated below.

9 External validation of the IPIM composite score

9.1 Some Expected Relationships

It will be clear that we expect a strong relationship between the variable political interest and the Interest in political issues in the media. Furthermore we expect that people with higher education will be more interested in political issues in the media. In principle we do not expect that age has an influence on interest in political issues in the media but we expect that education and age will have an effect on Political interest. So the relationship between age and IPIM is only indirect. These

expected relationships have been summarized in the path diagram presented in Figure 11.

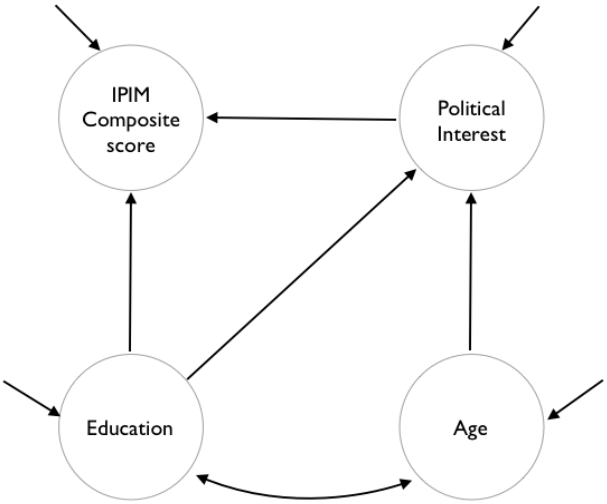

Figure 11: Theoretical model for the relationships of IPIM, political interest, age and education variables.

9.2 The model for validation

Table 5 shows the correlations between IPIM composite score and the variables political interest in ESS cumulative data. This table presents in the 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 5 contains in the second column for each country the correlations corrected for measurement error[20]. The disattenuated correlation ($\hat{\rho}_d$), corrected for measurement error, between the observed variables was computed as:

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

where xi and xi′ are the observed variables and Q_i and $Q_{i'}$ are the estimates of the qualities of the variables which were respectively .32 and .76 for IPIM and Political interest.

[20] The detailed overview of variables, quality coefficients and further analysis can be found from Appendix 5.

	Correlations between IPIM and 'political interest	
	without correction for measurement error	corrected for measurement error
Austria	.32	.64
Belgium	.33	.67
Bulgaria	.24	.49
Switzerland	.27	.55
Cyprus	.15	.31
Czech Republic	.31	.62
Germany	.35	.72
Denmark	.31	.62
Estonia	.23	.47
Spain	.23	.47
Finland	.26	.52
France	.29	.60
United Kingdom	.38	.77
Greece	.25	.51
Hungary	.27	.54
Ireland	.39	.78
Israel	.29	.59
Italy	.31	.63
Luxembourg	.30	.62
Netherlands	.33	.67
Norway	.29	.58
Poland	.32	.65
Portugal	.20	.40
Russian Federation	.24	.49
Sweden	.35	.70
Slovenia	.31	.63
Slovakia	.28	.57
Turkey	.13	.26
Ukraine	.27	.54

Table 5: Correlations between the IPIM Composite score and 'political interest'

The results show how large the difference is between the correlations corrected and not corrected for measurement error. After correction for measurement error the variable IPIM is highly correlated with 'political interest' 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.

Given this big effect of the measurement errors on the estimates of the effects the parameters of the model in Figure 10 are also estimated while correcting for measurement error. How this is done is illustrated in Appendix 5. In order to be sure that the model was correct the procedure of Saris, Satorra & Van der Veld (2009) was used to test for misspecifications in the model in the different countries. These corrections for misspecifications have led to the introduction of some effects,

which were not specified before. The results for the effects specified in the model of Figure 10 including the necessary correction have been presented in Table 6.

	PI→IPIM	EDU→IPIM	EDU→PI	AGE→PI	AGE→IPIM	AGE*EDU
Austria	.6	.15	.34	.21	ns	-.17
Belgium	.61	.18	.36	.11	ns	-.24
Bulgaria	.45	ns	.35	.33	.15	-.21
Switzerland	.51	.19	.29	.19	-.08	-.1
Cyprus	.18	.13	.3	.37	.47	-.48
Czech Republic	.58	.13	.3	.28	ns	-.11
Germany	.67	.15	.41	.27	-.06	-.14
Denmark	.55	.21	.28	.2	.12	-.16
Estonia	.45	.13	.29	.2	ns	-.19
Spain	.44	.2	.43	.08	.33	-.49
Finland	.44	.22	.32	.22	.25	-.37
France	.55	.17	.38	.21	ns	-.37
United Kingdom	.71	.16	.35	.21	.12	-.22
Greece	.47	.04	.33	.28	.36	-.52
Hungary	.49	.12	.36	.18	.14	-.2
Ireland	.7	.13	.32	.29	.23	-.3
Israel	.51	.12	.23	.22	.23	-.06
Italy	.54	.22	.48	.23	.3	-.4
Luxembourg	.52	.19	.34	.3	.15	-.24
Netherlands	.61	.18	.36	.15	.06	-.3
Norway	.46	.24	.34	.3	.22	-.21
Poland	.6	.11	.39	.28	.15	-.34
Portugal	.36	.14	.5	.14	.21	-.55
Russian Federation	.48	ns	.33	.25	.14	-.34
Sweden	.6	.23	.35	.23	.26	-.27
Slovenia	.58	.07	.32	.28	.16	-.27
Slovakia	.55	.04	.24	.21	.1	-.12
Turkey	.25	ns	.34	.12	.27	-.42
Ukraine	.51	ns	.25	.31	.16	-.29

Table 6: Coefficients for the model in the different countries. Abbreviations are PI (political interest), EDU (education), AGE (age), ns (not significantly different from 0).

These results completely support our first hypotheses because we found very strong effects from political interest on the variable IPIM for all countries, except Cyprus and Turkey.

The second hypothesis, about direct and positive effects from education to 'political Interest' and IPIM, found partial support. The effect of education on political interest is very clear and significant for all countries, but the direct effect on IPIM is weaker and in few cases not significant. This is due to the fact that education also has an effect on Political interest and due to that an indirect effect on IPIM, which can make the direct effect close to zero. However in most countries the standardized effect is significantly different from zero and substantially relevant.

Our third hypothesis is supported completely because significant effects between age and political interest were found for all 30 countries. And in addition, in

most cases there was also a significant link between age and IPIM. We expected this effect to be close to zero, however in several countries this effect was also significant and substantially relevant.

Figure 12 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 not significant.

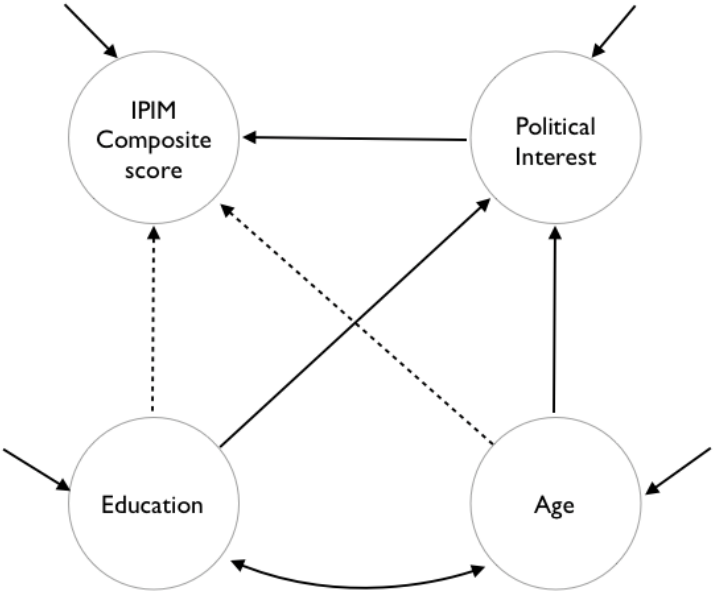

Figure 12: Final empirical model for most of ESS countries.

The strongest effects are from political interest on the IPIM composite score. This means that people who are more interested in politics spend more time for political issues in the media. However this was not the only effect on IPIM it turns out that the effects of education and Age direct on IPIM were often also significant which means that IPIM is not completely determined by Political interest. So it is found that IPIM is not exactly the same as Political interest other variables have also influence on IPIM.

Because the expected relationships have been found the external validity of the measure of IPIM has been confirmed.

10 Conclusions

This study of the measurement of IPIM (Interest in political issues in the media) has shown that the indicators for interest in political issues on the TV, the radio and the newspapers can be seen as scalar invariant reflective indicators for the concept. So the means of IPIM and the relationships of IPIM with other variables can be compared cross-countries and across time. This holds true for the means of the latent variable IPIM but also for the means of the composite scores for IPIM.

It was also shown that the composite score of IPIM has external validity because it has the expected relationships with other variables. In fact the relationship with Political interest is rather strong. However the relationship is not so strong that the two concepts cannot be distinguished. This was for example shown by the fact that Education and Age had next to an indirect effect also a direct effect on IPIM. This suggests that the effect on Political interest and IPIM are at least not exactly the same.

We have shown that it is not necessary to compute the composite score for IPIM by a weighted average because the un-weighted average of the indicators has approximately equal quality as the weighted composite score.

However it has been found that the quality of the Composite score is rather low in all countries and rather similar .32. This means that the composite score is only for 32% explained by the latent variable IPIM, the rest (68%) is error. For the comparison of the means of the composite score that is not very harmful, although they are different from the latent means. The means can still be compared. However comparing relationships between IPIM and other variables, based on the composite score, can not be done without correction for measurement error. The procedure for this has been explained in Appendix 5.

Appendices

Appendix 1: Methods for MTMM

These are the three different methods used for estimating the reliability and validity in order to decide which method was best for the ESS.

- Method used in the core questionnaire: "**8 numerical categories**".

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

- Method used in the complementary questionnaire: "**open numerical question**".

WRITE IN HOURS**:** ☐☐ AND MINUTES**:** ☐☐

- Method used in the complementary questionnaire: "**7 verbal categories**".

No time at all	01
Very little time	02
A little time	03
Some time	04
Quite a lot of time	05
A lot of time	06
A great deal of time	07

Appendix 2:

2.1.- Quantification of categories

The 8th category (more than 3 hours) cannot be directly determined because there is no upper boundary for this category. The solution we applied was to use information from the supplementary questionnaire in which respondents provided the total time spent on each of the media. The steps done to obtain a numerical value for the 8th category are detailed below:

- Respondents who mailed their questionnaire after the interview were deleted to avoid mode effects.

- Re-codification in the supplementary questionnaire: The original variable "test1a" from database expressed the total amount of hours watching television in a normal weekday, and "test1b" expressed the minutes. Therefore, in order to compute the total amount of minutes spent on television, the re-codification was $mnts_{TV} = hours_{TV} \times 60 + minutes_{TV}$. This was also done for *radio* and *newspaper*. After that, we removed 'impossible' values (larger than 960 minutes or 16 hours).

- We deleted respondents who chose more than 3 hours in the core questionnaire but did not report a response higher than 180 minutes in the supplementary questionnaire.

- Finally, we determined the value for the 8th category as the value that maximized the correlations between the transformed variable from the main questionnaire and the recodified minutes (*mnts)* variables from the supplementary questionnaire across the range of possible values (i.e. 180 to 960 minutes).

2.2.- Numerical values for countries in 2nd, 3rd and 4th ESS rounds

The countries not present in the first round of ESS and for this reason without additional information from SB-MTMM are Bulgaria, Cyprus, Estonia, Iceland, Latvia, Luxemburg, Romania, Russia, Slovakia, Turkey and Ukraine. The procedure followed for each country was to use transformed, numerical values from a country in the 1st ESS round. This country should have the most similar distribution for a medium in order to use its values for the new country. For choosing the similarity in distribution for a medium we check for the 'total time' and the 'time spent on political issues' for each medium. This means that each medium was selected independently. Even in this way, for almost all countries, the values chosen were from the same country for the different media, and moreover, this country is usually close geographically, with a few exceptions. In Figure A.2.2.1 we can see several

examples of the similarities in distribution between Austria and Luxemburg; between Poland and Slovakia and Ukraine; and between Cyprus and England.

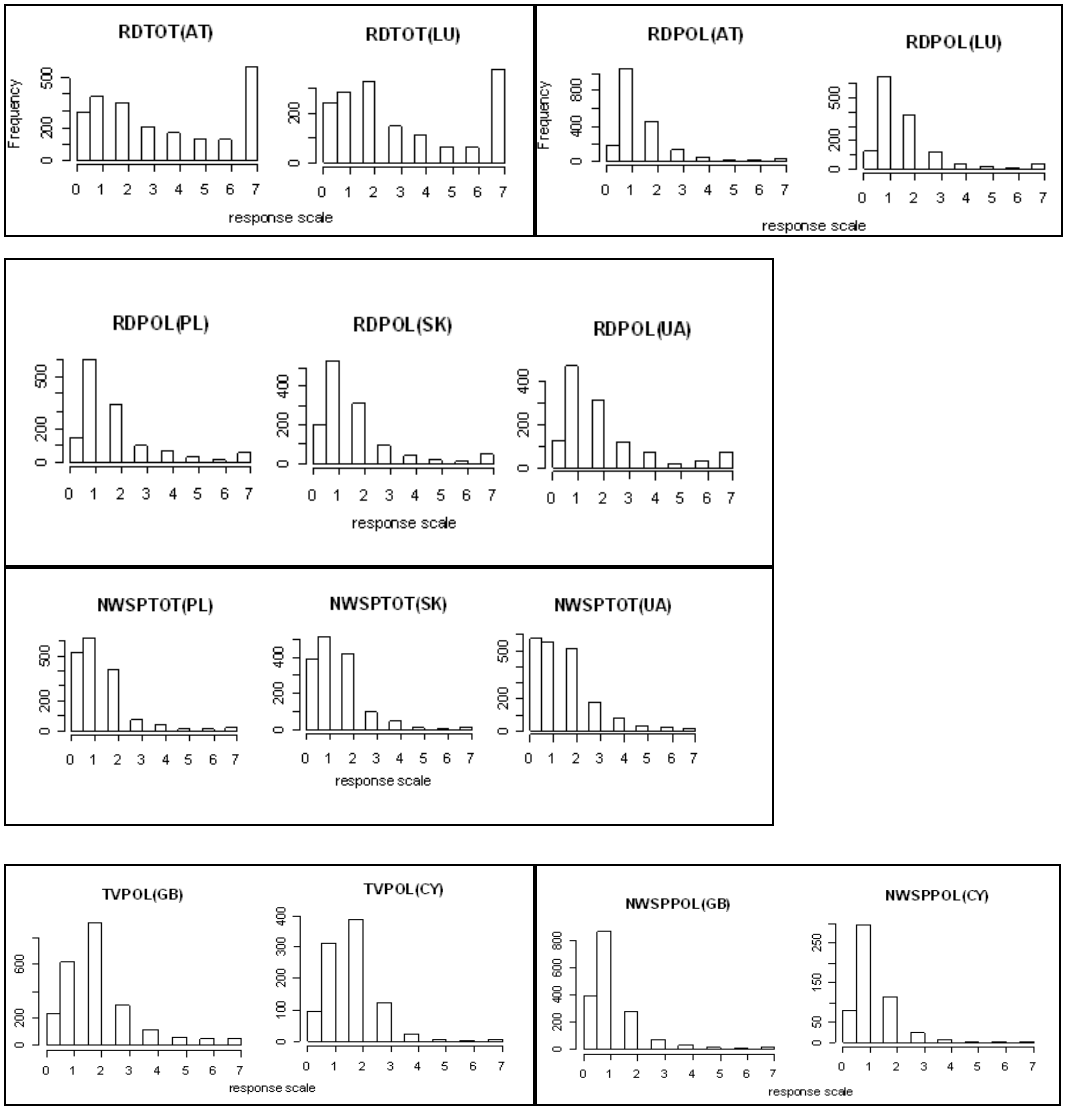

Figure A.2.2.1: Similar distributions with new countries

In the first comparison between Austria and Luxemburg the distribution is quite comparable for the three media, and for 'total time' and 'time spent in politics'. Also comparable are Poland, Slovakia and Ukraine for television, radio and newspaper. All these similar distributions are between countries quite close geographically. But this is not always the case, for instance Great Britain and Cyprus are also comparable as can be shown in the last histograms of the Figure A2.2.1.

For the remaining countries; Bulgaria has been given the values of Portugal for television and radio and from Poland for newspaper, Estonia got the values from the Czech Republic for television and radio and from Germany for newspaper; Iceland

received the values from Denmark for television and radio and from Germany for newspaper; Romania and Russia Federations got the scores from the Czech Republic for television and Portugal for radio and newspaper; Slovakia and Ukraine got values from Poland; and Turkey got values from Greece for all the media.

Appendix 3. SB-MTMM qualities

The qualities for the SB-MTMM experiment in the different countries in the ESS first round are shown in Table A.3:

	MTMM QUALITY		
	Television	**Radio**	**Newspaper**
Austria	.96	.85	.76
Belgium	.94	.96	.51
Switzerland	.56	.64	.92
Czech Republic	1.00	.66	1.00
Germany	1.00	.90	.74
Denmark	.92	.74	.88
Spain	.64	.85	.76
Finland	1.00	.88	.53
France	1.00	.58	1.00
Great Britain	.92	.71	.58
Greece	.87	1.00	.83
Ireland	.56	.81	1.00
Israel	.79	.79	.79
Italy	.76	.90	.76
The Netherlands	.98	.79	.96
Norway	1.00	.96	.62
Poland	1.00	.88	.67
Portugal	.62	.89	.73
Sweden	.69	.62	.79
Slovenia	.69	.77	.67

Table A.3: SB-MTMM qualities

Appendix 4: Definition and quality of the political variables

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 variables used are shown in Table A.4:

Variable	Label ESS database	Standardized CFA loadings
Placement on left right scale	*Lrscale**	---
Satisfaction with present state of economy in country	Stfeco	.789
Satisfaction with the national government	Stfgov	.788
Satisfaction with the way democracy works in country	Stfdem	.819
State of education in country nowadays	*Stfedu**	---
State of health services in country nowadays	*Stfhlth**	---
Government should reduce differences in income levels	Gincdif	.824
Gays and lesbians free to live life as they whish	freehms	.764
Ban political parties that wish overthrow democracy	prtyban	.845
Modern science can be relied on to solve environmental problems	scnsenv	.815

Table A.4: Questions for "political knowledge"

Using Confirmatory Factor Analysis (CFA) seven variables were selected which had the highest loading on the general factor, the final standardized factor loadings are shown in the last column of Table A.4. The variables omitted were Lrscale, Stfedu and Stfhlth. The variance of the general knowledge factor (F) was .622 while the composite score over the 7 items had a variance (CS) of .691. Based on this information we can define the quality (Saris & Gallhofer 2007a) of the knowledge variables as follows:

$$\text{Quality the knowledge indicator} = \frac{\text{var}(F)}{\text{var}(CS)} = \frac{.622}{.691} = .90$$

The error variance constrained in the model is computed as $(1-quality)\text{var}(\text{var}_i)$ for the different countries; where var(var_i) is the variance of the observable variable 'knowledge'.

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 Survey Quality Predictor (SQP[21]) and turned out to be .43. The same procedure $(1-quality)\text{var}(\text{var}_i)$ was used to estimate the error variance in each country, where var(var_i) is the variance of the observable variable 'discussion about politics'.

Political participation: This variable is operationalized by the sum of the variables ***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 still not available for other languages and we assume the quality of conventional and quality of protest actions as an average of the three countries (.59 for conventional actions and .65 for protest actions). In order to test the model, we will constrain the error variance of "political participation", which is formed by conventional and protest actions. We use the same formula as before:

$Var(\delta_i) = [1-Quality(_i)] * var(i)$ where i= conventional, protest actions

In this way, two error variances were obtained and the average of these two $var(\delta_i)$ was the error variance of the political participation, which was constrained in the model. This procedure was followed for each country.

[21] For more details about the Software *Survey Quality Predictor*, see http://www.sqp.nl/

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[22] 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:

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

Table A1. Questions for "political knowledge"

[22] Website: *www.sqp.nl*

Appendix 5: The definition and quality of variables

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[23] and turned out to be .76.

According to table 4 ***IPIM composite score (ipim_cs)*** qualities were rather similar between countries and across time, so for the purpose of simplicity, we used for all groups the same quality coefficient .32.

For the age and education we have used quality coefficient 1.

Overview of measurement procedure

Analysing relationships between these observed variables we used the model presented in Figure 10. As an example we have subsequently described LISREL inputs for Austria. First input is for estimation without correction for measurement error and second is with correction. The only difference is that for correction, we have on main diagonal for IPIM and political interest variables changed correlation 1 with specific quality coefficients (.32 and .76).

LISREL input for estimation of the Austria model without correction for measurement error:
Analysis of the sum scores without correction for measurement error
data ni=4 no=1750 ma=km
km
*
1
.32 1
-.318 .76 1
.066 -.129 1
.189 -.269 -.173 1
la
ipim_cs polintr age edu
mo ny=2 nx=2 ps=sy,fi ga=fu,fi be=fu,fi
fr ga 1 2 ga 2 2 ga 2 1
fr be 1 2
fr ps 1 1 ps 2 2
pd
ou

[23] Website: www.sqp.nl

LISREL input for estimation of the model with Austria correction for measurement error:
Analysis of the sum scores with correction for measurement error

```
data  ni=4 no=1750 ma=km
cm
*
.32     ! The quality estimate for IPIM index
-.318  .76   !  The quality estimate for political interest
.066 -.129 1
.189 -.269 -.173 1
la
ipim_cs polintr age edu
mo ny=2  nx=2  ps=sy,fi ga=fu,fi be=fu,fi
fr  ga 1 2 ga 2 2 ga 2 1
fr be 1 2
fr ps 1 1 ps 2 2
pd
ou
```

References

Aarts, K., and Semetko, H.A. (2003). The divided electorate: Media use and political involvement. *The Journal of Politics, 65*(3), 759-784.

Andrews, F.M. (1984). Construct validity and error components of survey measures: a structural modeling approach. *Public Opinion Quarterly*, 48, 409-442.

Bollen, K.A. (1989). *Structural Equations with Latent Variables*. Willey

Byrne, B. M., R. J. Shavelson, and Muthén, B. (1989). "Testing for the equivalence of factor covariance and mean structures: the issue of partial measurement invariance." Psychological Bulletin 105(3): 456-466.

Campbell, D.T., and Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81–105.

Coromina Ll. and W.Saris (2009) Quality of media use measurement. *International Journal of Public Opinion* .

Coromina, L., Saris, W.E., and Oberski, D. (2008) "The Quality of the measurement of interest in the political issues presented in the media in the ESS", *Ask, Research & Methods*, 17; pp.7-30.

Eveland, W.P., & Scheufele, D. (2000). Connecting News Media Use with Gaps in Knowledge and Participation. *Political Communication, 17*, 215-237.

Guillén, L., Coromina, L., and Saris, W.E. (forthcoming): Measurement of Social Participation and its Place in Social Capital Theory

Guo, Z. (2000). Media use habits, audience expectations and media effects in Hong Kong's first legislative council election. *Gazzette, 62*(2), 133-151.

Harkness, J., F. van de Vijver, and Mohler, P. (2007). Cross-Cultural Survey Methods. Wiley, New York

Horn, J. L., J. McArdle, and Mason, R. (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, R. J. A., and Rubin. D.B. (1987): Statistical Analysis with Missing Data. Wiley, New York.

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

McLeod, J. M., Daily, K., Guo, Z., Eveland, W. P., Bayer, J., Yang, S., and Wang, H (1996). Community integration, local media use, and democratic process. *Communication Research, 23*, 179-209.

McLeod, J. M., Scheufele, D. A., & Moy, P. (1999). Community, Communication and Participation: The Role of Mass Media and Interpersonal Discussion in Local Political Participation. *Political Communication, 16*, 315-336.

Meredith, W. (1993). "Measurement invariance, factor analysis and factorial invariance." Psychometrika **58**: 525-543.

Newton, K. (1999). Mass media Effects: Mobilization or media malaise? *B. J. Pol., 29*, 577-599.

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

Oberski, D. and Saris, W.E. (2007). Differences in quality of questions across countries. In Festschrift for Jaak Billiet, University of Leuven.

Oberski, D. W.E. Saris and Hagenaars J.A. (2010). "Categorisation errors and differences in the quality of questions across countries." In J.Harkness et. al. (Eds) *Multinational and Multiregional Survey Methods*. Hoboken, NJ, Wiley, pp 435-453.

Potter, W.J. (1997). The Problem of Indexing Risk of Viewing Television Aggression. *Critical Studies in Mass Communication*, 14, 228-248

Saris, W.E., and Andrews, F.M. (1991): Evaluation of measurement instruments using a structural modeling approach. In Biemer, PP., Groves, R.M., Lyberg, L.E., Mathiowetz, N.A., and Sudman, S. (Eds.): *Measurement Errors in Surveys*. Wiley, New York, 575–597.

Saris, W.E, and Gallhofer, I. (2007a). *Design, Evaluation, and Analysis of Questionnaires for Survey Research*. Wiley

Saris, W.E, and Gallhofer, I. (2007b). The results of the MTMM experiments in round 2. ESS Report

Saris, W.E., Satorra, A., and Coenders, G. (2004). A New Approach to Evaluating the Quality of Measurement Instruments: The Split-Ballot MTMM Design. *Sociological Methodology*, 34, 311-347.

Saris, W. E., A. Satorra, and Sörbom, D. (1987). Detection and correction of structural equation models. *Sociological Methodology,* 17: 105-131.

Saris, W.E., Satorra, A., and van der Veld, W. (2009). Testing Structural Equation model or Detections of misspecifications? *Structural Equation Modeling,* 16, 561-582.

Saris, W. E., A. Satorra, and Sörbom, D. (1987). "Detection and correction of structural equation models." Sociological Methodology **17**: 105-131.

Saris, W.E., A.Satorra and W.Van der Veld 2009) Testing Structural equation models or detection of misspecifications? Structural Equation Modeling,16: 561-582.

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

Slater, M.D. (2004). Operationalizing and analyzing exposure: The foundation of media effects research. *Journalism and Mass Communication Quarterly*, 81(1), 168-183.

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

Swyngedown, M., and Goeminne, B. (2007). On the relationship between media use, social capital, social-political attitudes and voting behaviour. In G. Loosveldt, M. Swyngedown and B. Cambré (Eds.), *Measuring Meaningful data in social research* (pp. 113-140). Leuven: Acco.

Steenkamp, J. and Baumgartner, H. (1998). "Assessing measurement invariance in cross-national consumer research." Journal of Consumer Research **25**: 78-90.

Little, R. J. A. and D. B. Rubin (1987) Statistical Analysis with Missing Data. New York: John Wiley.

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

Vreese, C.H.d., and Semetko, H. A. (2004). News matters: Influences on the vote in the Danish 2000 euro referendum campaign. *European Journal of Communication, 43*, 699-722.

[i] An SPSS file with all data for all respondents of round 1 to 4 can be found in *http://www.upf.edu/survey/working/extrafiles*

[ii] This part has been based on a paper of Coromina and Saris (2009)

[iii] The part is a considerably adjusted version of a paper of Coromina, Saris and Oberski (2008)