Political Knowledge in Comparative Perspective: The Problem of Cross-National Equivalence of Measurement

Martin Elff*

Prepared for delivery at the MPSA 2009 Annual National Conference, April 2–5, 2009, Palmer House Hilton, Chicago, Il.

Abstract

The Comparative Study of Electoral Systems (CSES) presents a unique resource for comparative research on political attitudes and behavior. From the beginning, country components of the CSES have contained each at least three items concerned with political information and knowledge. These items vary considerably across countries in terms of question format and question content. Using methods of Item Response Theory (IRT), the paper examines how these aspects impinge on the discriminance and difficulty of the items, both important aspects of their validity as indicators of political knowledge. It shows that the question content is especially important for the items difficulty: Notwithstanding the political context, items that ask for numbers (e.g. of federal states or EU member countries) are much more difficult to answer, given the level of political knowledge. Further, notwithstanding the political context, questions about foreign policy matters have a higher discriminance, that is, can better distinguish between different levels of political knowledge, than other items. The paper concludes with a discussion on how cross-national equivalence of knowledge questions can be enhanced.

That the mean level of political information in the American mass public is low and its variance is high—at least as from the point of view of civics texts and classical political theory—is one of the best documented results of empirical political science (Berelson *et al.* 1954; Campbell *et al.* 1960; Neuman 1986; Converse 1990, 2000). There seems hardly any reason to doubt that the situation is different in other countries. Nevertheless, one may

* University of Mannheim, Department of Social Sciences, 68131 Mannheim, Germany, e-mail: elff@sowi.uni-mannheim.de

hypothesize that levels of political information or the consequences of inequality in political information may differ across countries. It has been argued in the literature that heuristics may serve as a remedy for a lack of political information, but also that the efficiency of heuristics will depend on features of the context (Kuklinski *et al.* 2001; Jerit *et al.* 2006; Lau and Redlawsk 2008), such as the social environment (Huckfeldt and Sprague 1995), the media system (Prior 2007), or the incentives provided by the institutional structure (Lupia and McCubbins 1998). If one would intend to devise policies to ameliorate the consequences of the unequal distribution of political information and thereby to promote political equality, it may be crucial to be able to analyze the implications of existing institutional arrangements and social structures in different countries. In order to achieve this, a minimum requirement is to be able to measure political information in a cross-nationally equivalent way. Otherwise, it would be futile to make any statements about the resource-dependence of political knowledge (Gordon and Segura 1997; Berggren 2001; Grönlund and Milner 2006) or about the consequences of institutions for political equality, in so far as it is affected by different levels of political knowledge (Verba 2003).

The Comparative Study of Electoral Systems (CSES) presents one of the rare opportunities for a comparative assessment of political knowledge of citizens from different countries, and for its impact on political choices in a comparative perspective. Each national electoral study that contributes to the CSES includes three questions on political knowledge in its questionnaire (CSES 2007). While are cross-national studies on political knowledge are rare (some are mentioned in Delli Carpini and Keeter 1996: 90) only the CSES makes data available on political knowledge in 36 countries for roughly the same period of time (that is, the first decade of the new millennium). For this reason, it is of considerable importance for comparative studies of political knowledge whether measures of political knowledge derived from the items contained in these studies are comparable across countries.

Most of the literature that deals with the problem of comparability of measures, or their *equivalence*, is concerned with the consequences of a divergence in the meaning of cultural phenomena across countries, including the meaning of terms used in survey questionnaires (Anderson 1967; Przeworski and Teune 1966; King *et al.* 2004). However, the question about the cross-national equivalence of political knowledge measured based on CSES data is much more pressing. The political knowledge items used in the CSES participant countries are not just more or less successful translations of a standard set of items. On the contrary, they vary considerably in terms of content and format. The consequences of these variations is the subject of this paper. It examines how question topic, item format and the type of answer

required by the knowledge questions affect the quality of measurement, and consequently, the comparability of knowledge measures across countries.

One could argue, of course, that measuring political knowledge is sufficiently straightforward so that the question of equivalence does not arise, or, if it does, only in a relatively mild form: If respondents are given "quiz items" about political matters, then almost by definition one does not measure anything but political knowledge. If this assumption were true, than all that would be needed for the comparability of scores constructed based on those knowledge questions is that the distribution of the scores is the same in all countries. As the next section shows, however, the distribution of knowledge scores varies considerably across countries. If cross-country differences in the distribution of knowledge scores were only a consequence of variations in the difficulty of knowledge questions, one could consider rescaling the knowledge scores so that they have the same mean and same variance in all countries (after stratification where appropriate). However, such an adjustment presupposes that all knowledge questions are equivalent in how well they tap political knowledge. A second section therefore is used to examine how well knowledge questions discriminate between different levels of political knowledge. Unfortunately, the results of that question indicate a considerable cross-country variation in discrimination of levels of political knowledge by the items. Even worse, some of the knowledge questions seem ill-suited to measure general political knowledge. The dismal evidence of the second section begs the question about the causes of the variation in the performance of knowledge question used in the CSES participant countries. A systematic examination of the impact of question topic and format therefore is given in the third section. The paper then concludes by daring some suggestions on how the measurement of political knowledge can be improved.

1 Some first evidence of equivalence problems

Cross-country comparability of metric measures for political knowledge means, first and foremost, that identical values on a political knowledge scale indicate more or less equivalent levels of political knowledge. If such a scale is constructed on a limited number of questions answered correctly—three in the case of the CSES—then the measurement of knowledge is accordingly coarse. Comparability can then only mean that individuals who obtain the same value on this scale lie in the same range of political knowledge. Assuming that the distribution of political knowledge is the same in all countries under study (certainly an assumption that is debatable), a minimal empirical criterion for the comparability of such coarse knowledge measures is that in all countries the proportion of respondents that answer

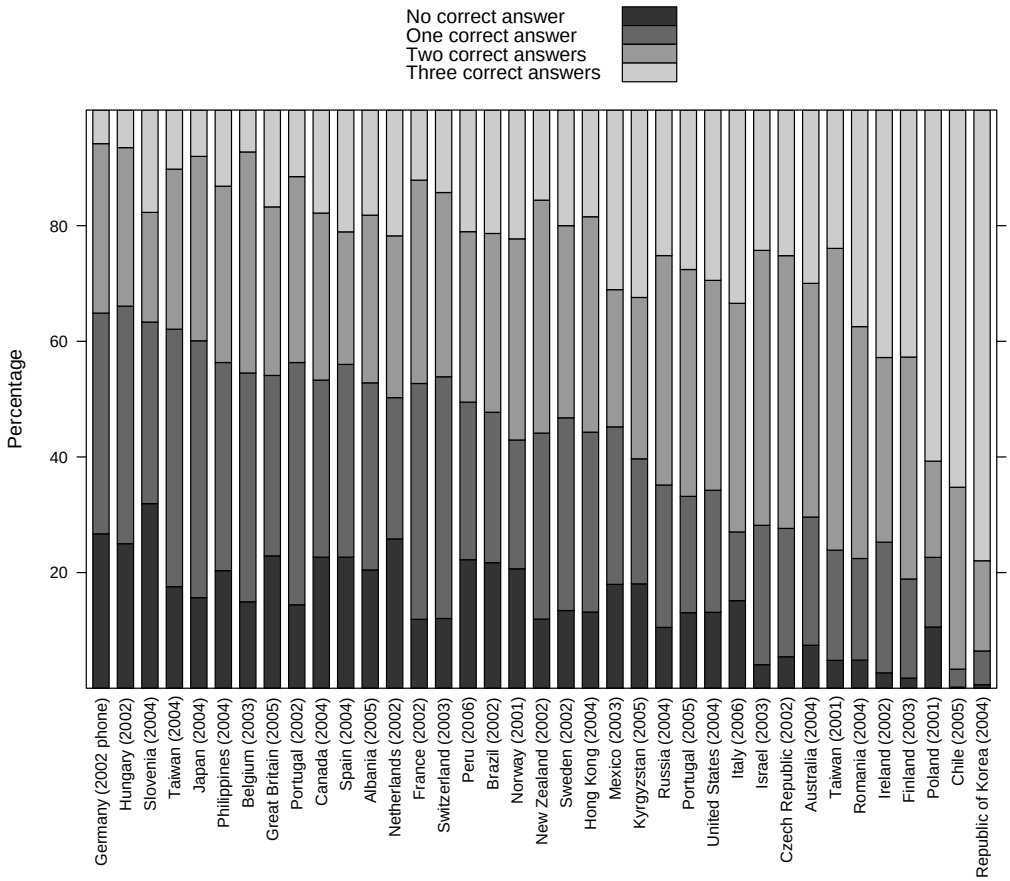
one, two, or three questions correctly, is, barring sampling error, essentially the same. In case of the CSES knowledge scales the evidence suggests otherwise. Figure 1 shows a barchart of the percentages of respondents answering three, two, one, or none of the political knowledge questions correctly. The individual stacks of bars correspond to the country samples of CSES, Module 2. They are sorted in increasing order by the mean number of correct answers. The figure clearly shows that the proportion of respondents correctly answering one, two, or three questions varies considerably across country samples, but in one case also between two samples from the same country.

Certainly, such cross-country variations in the rates of correct responses may be the result of variations in the distribution of political knowledge, as a consequence of variations in the average educational attainment of the populations of these countries (assuming that education is a powerful predictor of a cognitive ability such as political knowledge). For this reason, Figure 2 depicts the proportions of three, two, one or no correct answers to the three CSES political knowledge questions stratified by education. Stratified percentages are computed as follows: As a first step the percentages three, two, one, or no correct answers are computed for each level of education in each country. In a second step, the means of these percentages are computed for each country across educational strata.[1]

The figure reveals that there are quite considerable differences between the country samples in terms of the numbers of correctly answered knowledge questions. On the one extreme, in Slovenia (2004) more than a quarter of the respondents, stratified by education, were unable to answer any of the three knowledge questions correctly. On the other extreme, in Chile (2005) and the Republic of Korea (2004) three out of five respondents were able to answer all three knowledge questions correctly. It seems difficult to discover any pattern in the ranking of the country samples in terms of the percentages of correct answers: For example, Asian countries (e.g. Republic of Korea) have rankings next to European countries (e.g. Finland), new democracies (such as Kyrgyzstan) have rankings next to very old and established ones (such as the United States). Further, Taiwan in 2001 has a very high ranking in measured political knowledge (9th rank), while in 2004 Taiwan obtains a very low rank (6th last ranking place). It does not seem very plausible that the "performance" of the respondents in the different countries reflect real differences between the respective populations. Rather it is likely that these differences reflect differences between participating

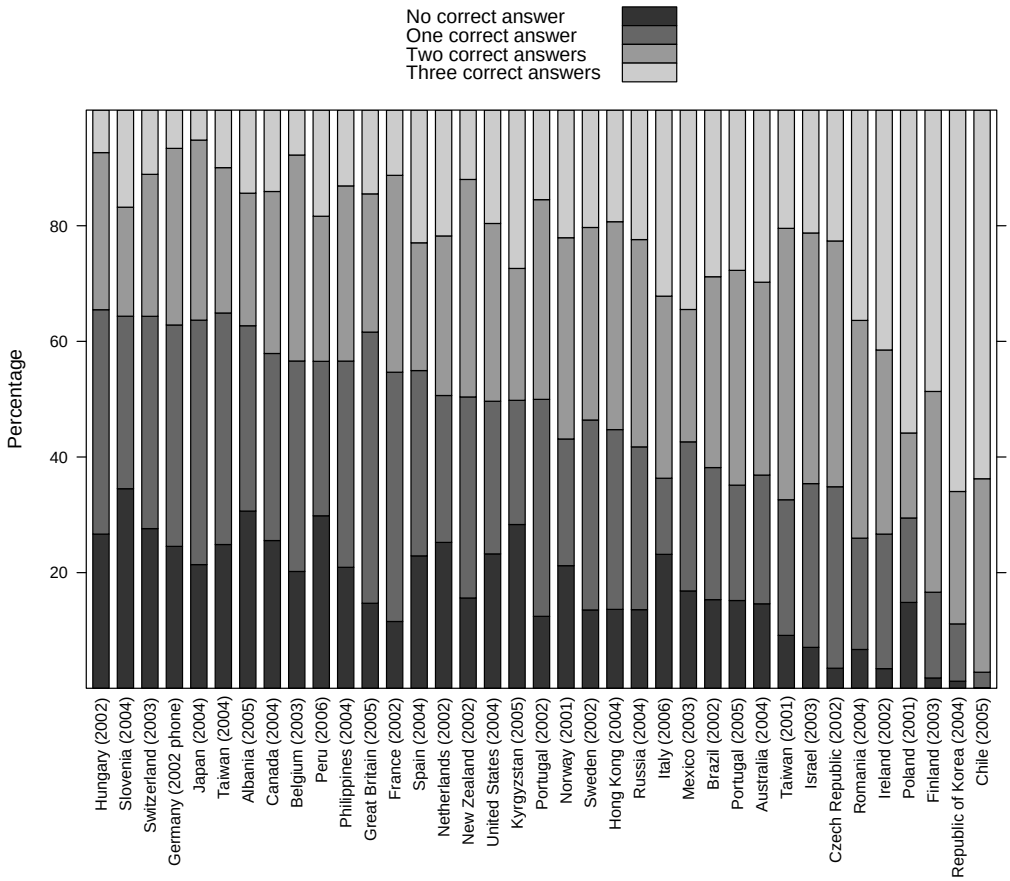
[1]The number of educational categories used varies across country samples. In order to achieve at least a rough comparability of educational strata, educational levels are collapsed into four categories such that the limits between adjacent categories are the same in all countries. How these collapsed categories correspond to the original educational categories in the country samples is summarized in the appendix of this paper. The appendix also explains the rationale behind the construction of stratified percentages.

Figure 1: Responses to knowledge questions used in the individual components of the *Comparative Study of Electoral Systems*, Module 2 (CSES2).



Source: Comparative Study of Electoral Systems (CSES) – Module 2 (2001-2006), Full Release – June 27, 2007.
Note: No knowledge questions were asked in the Bulgaria (2001), Germany (2002) mailback, Denmark (2001), and Iceland (2003) survey components.

Figure 2: Responses to knowledge questions used in the individual components of the *Comparative Study of Electoral Systems*, Module 2 (CSES2)—stratified by education.



Source: Comparative Study of Electoral Systems (CSES) – Module 2 (2001-2006), Full Release – June 27, 2007.
Note: No knowledge questions were asked in the Bulgaria (2001), Germany (2002) mailback, Denmark (2001), and Iceland (2003) survey components.

studies in terms of measuring knowledge questions. Of course, cross-country differences in terms of rates of correct answers that seem to defy systematic explanation can only give a hint that there may be problems in the measurement of knowledge or with regards to the equivalence of knowledge measures. A more systematic study of the performance of knowledge questions in the CSES is presented in the next section.

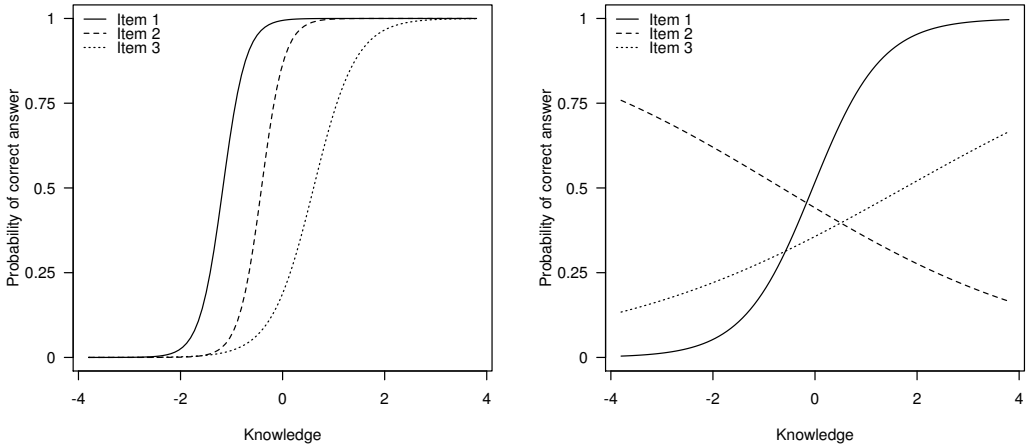
2 Performance of Questions about Factual Knowledge in CSES Country Components

Prima facie the measurement of political information seems, compared to constructs like political trust, party identification, political interest or political efficacy, relatively straightforward. In contrast to political attitudes, the measurement of such a cognitive ability is unlikely so suffer from problems such as social desirability, halo effects, and acquiescence. Nevertheless, the measurement of a specific ability by letting respondents perform cognitive tasks may have its own kind of challenges (Luskin 1987; Kuklinski and Quirk 2001; Mondak 2000, 2001; Prior and Lupia 2008; Sturgis *et al.* 2008). Tasks need to be selected such that they tap the ability one intends to measure and that the performance on these tasks is more dependent on this ability than on other ones. That is, one needs to make sure that the tasks have an appropriate level of *discriminance* for the ability in question. The discriminance of a cognitive task may suffer from a lack of validity, if e.g. the performance on a political information item depends more on the verbal ability of respondents or their ability to calculate or to recall numeric data than on their political awareness; or from a lack of reliability, e.g. if an item is ambiguous or allows respondents to guess the correct answer. Political knowledge items in "multiple choice" format, that is, that present respondents with a set of given answers from which to choose, may be more vulnerable to guessing than knowledge items that ask a question in "open" format.

An appropriate set of political information items also needs to have some variation in terms of *difficulty*. Suppose a political knowledge item has perfect discriminance, that is, one could be completely confident that a respondent answering the corresponding question correctly has at least a specific level of political knowledge. In that case, additional items of the same difficulty will add nothing to the individuals' level of political knowledge, unless they are more difficult than the first item.[2]

[2] Additional items of the same difficulty may, however, add to the posterior probability that a respondent surpasses a certain level of political information if the discriminance of the first item is less then perfect.

Figure 3: Item characteristic curves for knowledge questions in two participant countries in the CSES2.



(a) United States (2004) (b) Japan (2004)

Source: Comparative Study of Electoral Systems (CSES) – Module 2 (2001-2006), full release – June 27, 2007.
Note: The item characteristic curves are derived from model estimates presented in Table 1.

The appropriate approach to assessing the performance of knowledge measures based on binary items is logistic latent-trait modelling (LTM), one element of the repertoire of Item-Response-Theory (IRT) (McDonald 1999; Bartholomew *et al.* 2008; van der Linden and Hambleton 1997). In the logistic latent trait model, the log-odds of an individual i to give a correct answer to a knowledge question j (or more generally, to successfully complete a task in a cognitive ability test) is assumed to be a function of her or his (latent) knowledge (or cognitive ability), the discriminance of the item, and its difficulty. In its classical form, this function is given by

$$\log \frac{\Pr(X_{ij} = 1|U_i)}{\Pr(X_{ij} = 0|U_i)} = \beta_j(U_i - \alpha_j) \tag{1}$$

where X_{ij} is an observed random variable that takes the value one if individual i answers correctly to item j and the value zero otherwise. The discriminance of the item is given by the slope parameter β_j and its difficulty is given by the shift parameter α_j.

Figure 3 gives two examples on how knowledge questions may relate to political knowledge as a latent variable. The left panel (Figure 4(a)) shows the item characteristic curves of the knowledge items contributed by the American National Election Study 2004 to CSES

Module 2. The item characteristic curves in the plot indicate how the probability that an individual answers a knowledge question correctly is related to her or his political knowledge. The curves are quite steep, indicating a strong relation between the probability of a correct answer and political knowledge. That is, the items discriminate well between different levels of political knowledge and differ in difficulty: If an individual answers at least one, two, or three knowledge questions correctly, one can be fairly certain that he or she achieves at least a specific minimum level of political knowledge. The right panel (Figure 4(b)) shows how measurement of political knowledge can go wrong: The item characteristic curves of item 3 is very flat, answering the corresponding question seems only weakly related to political knowledge. Even worse, the item characteristic curve of item 2 indicates that the probability to correctly answer the corresponding question *decreases* with political knowledge. Consequently, the number of knowledge questions correctly answered can hardly serve as an indicator of political knowledge in this case.[3]

Table 1 reports these estimates for each of the three items in each of the country samples of Module 2 of the CSES. The estimates are computed on the base of marginal maximum likelihood. Details about this estimation procedures are given in the appendix of this paper.[4] For each country sample, the table reports the difficulty and discriminance parameter estimates for each of the three items in the participating electoral study. For each country, the items are ordered in such a way that "item 1" denotes the easiest knowledge item in the respective electoral study, that is, the one with the highest relative number of correct answers, "item 2" denotes the second easiest or second difficult one, and "item 3" denotes the most difficult knowledge item of the respective study.[5]

[3] Krosnik *et al.* (2008) discuss some problems pertaining to the reliability of a knowledge scale derived from the 2004 ANES items: In contrast to earlier ANES, it was the interviewers task to code answers as correct or incorrect, in which they seem to have exercised some "leniency" in favor of the respondents, for example, by coding the (slightly) incorrect response "is prime minister of England" to the question about Tony Blair's current job, as correct. These problems surfaced after an examination of the ANES data in reaction to problems connected to the item involving William Rehnquist, the Chief Justice of the Supreme Court, indicated by Gibson and Caldeira (2008). These problems seem to become apparent in the fact that item 2, the question about Tony Blair, has a slightly decreased difficulty, indicated by the position of its item characteristic curve in Figure 3, and by the somewhat flatter slope of the ICC of item 3, the question about William Rehnquist. This lead Krosnik *et al.* (2008) to add a warning "to use the political knowledge questions only with great caution" to the ANES documentation. Compared to the problems that involve e.g. the Japanese knowledge questions, the ANES problems seem relatively benign.

[4] Unfortunately, the software used for computing the estimates does not allow for weighted data. Therefore, the IRT models are fitted to unweighted and unstratified data.

[5] The numbering and ordering of the items in the table and in the analyses troughout refer only to those knowledge items contributed to the CSES. At least the American National Election Study of 2004, however, uses more one more knowledge question, which is not included in the CSES data set.

Table 1: Item constants and discriminances of knowledge questions in CSES2 survey components; IRT-model estimates based on marginal maximum-likehood.

Sample	Item 1				Item 2				Item 3				Log-Likelihood	N
	Difficulty		Discrim.		Difficulty		Discrim.		Difficulty		Discrim.			
Albania (2005)	−0.89	(0.08)	1.83	(0.25)	0.11	(0.05)	2.00	(0.30)	0.88	(0.08)	2.25	(0.38)	−1853.6	1116
Australia (2004)	−1.39	(0.18)	1.44	(0.31)	−0.64	(0.09)	1.52	(0.36)	0.48	(0.11)	0.61	(0.10)	−2984.1	1769
Belgium (2003)	−1.08	(0.14)	2.39	(0.87)	0.28	(0.06)	1.24	(0.27)	3.11	(0.50)	0.63	(0.11)	−3408.4	2225
Brazil (2002)	−0.63	(0.05)	1.56	(0.14)	−0.09	(0.03)	2.68	(0.41)	0.78	(0.06)	1.30	(0.11)	−4517.9	2514
Canada (2004)	−0.56	(0.06)	1.77	(0.24)	0.18	(0.06)	1.03	(0.11)	0.67	(0.06)	2.23	(0.37)	−3078.9	1674
Chile (2005)	−2.04	(0.15)	1.83	(0.24)	−1.18	(0.27)	12.47	(21.15)	−1.91	(1.27)	0.38	(0.24)	−1041.2	1200
Czech Republic (2002)	−3.32	(1.06)	0.49	(0.17)	−0.57	(0.09)	4.69	(9.23)	1.63	(0.70)	0.28	(0.11)	−1631.3	948
Finland (2003)	−2.71	(0.49)	1.21	(0.31)	−1.59	(0.30)	0.86	(0.20)	−0.06	(0.05)	1.69	(0.61)	−1723.5	1196
France (2002)	−3.84	(1.35)	0.30	(0.11)	0.19	(0.14)	5.87	(9.53)	2.54	(0.76)	0.34	(0.10)	−1839.6	1000
Germany (2002 phone)	−0.33	(0.07)	0.94	(0.16)	0.17	(0.05)	1.19	(0.23)	2.19	(0.27)	1.12	(0.20)	−3401.9	2000
Great Britain (2005)	−0.39	(0.08)	1.25	(0.22)	−0.14	(0.06)	2.14	(0.55)	1.10	(0.14)	1.25	(0.21)	−1540.7	860
Hong Kong (2004)	−0.87	(0.33)	1.27	(0.76)	0.42	(0.31)	0.34	(0.15)	0.26	(0.13)	1.16	(0.67)	−1128.9	582
Hungary (2002)	−0.55	(0.08)	2.39	(0.85)	0.89	(0.15)	0.78	(0.13)	1.70	(0.21)	1.55	(0.33)	−1935.7	1200
Ireland (2002)	−2.92	(0.26)	1.39	(0.18)	−0.52	(0.08)	7.97	(6.80)	−0.02	(0.04)	1.29	(0.12)	−3269.4	2367
Israel (2003)	−2.17	(0.32)	1.74	(0.49)	−0.61	(0.11)	1.67	(0.57)	1.16	(0.23)	0.74	(0.17)	−1772.1	1212
Italy (2006)	−0.74	(0.20)	12.23	(37.25)	−0.66	(0.04)	4.80	(1.00)	0.57	(0.09)	0.83	(0.09)	−2203	1439
Japan (2004)	−0.05	(0.06)	1.48	(1.55)	−0.65	(0.40)	−0.36	(0.22)	1.75	(1.03)	0.34	(0.21)	−2927.9	1977
Kyrgyzstan (2005)	−0.88	(0.05)	2.31	(0.21)	−0.32	(0.03)	3.04	(0.36)	0.38	(0.04)	2.45	(0.25)	−3239.4	2000
Mexico (2003)	−0.95	(0.08)	1.02	(0.09)	−0.07	(0.04)	2.13	(0.26)	0.08	(0.03)	3.19	(0.63)	−3600.9	1991
Netherlands (2002)	−0.44	(0.04)	2.29	(0.26)	−0.11	(0.04)	2.36	(0.28)	0.77	(0.06)	1.91	(0.19)	−2750.4	1574
New Zealand (2002)	−2.60	(1.40)	0.32	(0.18)	−0.08	(0.06)	1.36	(1.23)	1.36	(0.73)	0.41	(0.23)	−3126.6	1741
Norway (2001)	−0.53	(0.04)	2.48	(0.30)	−0.48	(0.04)	2.13	(0.22)	0.85	(0.06)	1.75	(0.17)	−3476.7	2052
Peru (2006)	−0.56	(0.03)	4.43	(1.39)	0.31	(0.04)	1.56	(0.15)	0.59	(0.06)	1.22	(0.11)	−3590.3	2032
Philippines (2004)	−0.15	(0.10)	0.69	(0.16)	−0.14	(0.09)	0.72	(0.17)	0.78	(0.15)	1.25	(0.36)	−2384.9	1200
Poland (2001)	−0.90	(0.05)	2.59	(0.24)	−0.81	(0.04)	2.96	(0.32)	−0.79	(0.04)	2.78	(0.28)	−2514.1	1794
Portugal (2002)	−1.31	(0.15)	1.66	(0.36)	0.34	(0.06)	1.89	(0.48)	1.69	(0.21)	1.03	(0.17)	−2047	1303
Portugal (2005)	−0.82	(0.05)	2.66	(0.36)	−0.84	(0.05)	1.80	(0.17)	0.70	(0.05)	1.48	(0.13)	−4505.5	2801
Republic of Korea (2004)	−3.56	(0.54)	1.39	(0.31)	−1.25	(0.13)	8.25	(18.56)	−1.59	(0.16)	1.42	(0.23)	−1222.4	1500
Romania (2004)	−1.93	(0.17)	1.75	(0.28)	−1.04	(0.10)	1.59	(0.27)	0.33	(0.06)	1.79	(0.36)	−2162.4	1913
Russia (2004)	−1.56	(0.13)	1.34	(0.16)	−0.52	(0.05)	3.30	(0.95)	0.76	(0.07)	1.96	(0.32)	−2308.9	1496
Slovenia (2004)	−0.33	(0.06)	1.61	(0.20)	0.51	(0.06)	2.26	(0.33)	0.72	(0.06)	2.69	(0.46)	−1686.6	1002
Spain (2004)	−0.85	(0.08)	1.39	(0.16)	0.26	(0.04)	3.72	(1.06)	0.67	(0.07)	1.84	(0.24)	−2065.1	1212
Sweden (2002)	−1.27	(0.16)	1.40	(0.28)	0.11	(0.06)	1.54	(0.33)	0.77	(0.12)	1.09	(0.19)	−1707.3	1060
Switzerland (2003)	−1.74	(0.26)	1.04	(0.21)	0.83	(0.14)	0.84	(0.16)	0.72	(0.11)	1.63	(0.44)	−2407.7	1418
Taiwan (2001)	−1.86	(0.12)	2.86	(0.59)	−0.65	(0.04)	5.62	(5.99)	1.42	(0.18)	0.76	(0.11)	−2648.2	2022
Taiwan (2004)	−1.13	(0.10)	1.95	(0.34)	0.65	(0.06)	2.85	(0.86)	1.64	(0.17)	0.93	(0.12)	−2803	1823
United States (2004)	−1.17	(0.07)	4.41	(1.17)	−0.41	(0.05)	4.50	(1.43)	0.61	(0.06)	2.41	(0.39)	−1521.6	1066

Note: Standard errors in parentheses. Estimates are computed using R-package *ltm* (Rizopoulos 2008; R Development Core Team 2008) with unweighted data.
Source: Comparative Study of Electoral Systems (CSES) – Module 2 (2001-2006), full release – June 27, 2007.

As one may already suspect from the results reported in the previous section, item difficulties vary considerably across countries and across items. Of course, moderate differences between the item difficulties, which may become apparent in the item intercepts, are not a problem. Rather, a good cognitive test will employ tasks with varying difficulty, such that mastering a more difficult task is indicative of a higher degree of ability. If one compares the difficulty estimates in Table 1 one indeed finds that in most of the countries the items seem to exhibit this desirable property. Yet a comparison of the item difficulties along the columns lead to doubts about the equivalence of scores based on the number of correct answers to knowledge questions: The difficulties of the easiest, medium difficulty and most difficult items varies considerably across countries. For example, the difficulty of item 1 in the Korean electoral study of 2005 is estimated as -3.56 (extremely easy), whereas the corresponding estimate for the Japanese electoral study of 2004 is -0.05. The estimate of 3.11 for item 3 in the Belgian electoral study of 2003 indicates a extremely difficult knowledge question, whereas an estimate of -1.91 indicates that the supposedly most difficult knowledge question used in the Chilean electoral study of 2005 is indeed a very easy one.

The results with regard to the item difficulties are not the only reasons to become doubtful about the equivalence of knowledge measures used in the CSES Module 2 components. Item discriminances, that is, the degree to which the items reflect a common dimension of political knowledge, also vary considerably. One extreme is a discriminance coefficient of 12.47 estimated for item 2 in the Chilean electoral study, which would suggest that this item has an almost prefect relation to the underlying latent variable of political knowledge. Yet with 21.15 the standard error is even higher than this estimate. That is, although this estimate has a very large value, it is nevertheless statistically indistinguishable from zero, which is likely a side-effect of the relatively poor performance of item 3 (the discriminance coefficient of which is estimated as 0.38). Another extreme is the *negative* estimate of the discriminance of item 2 in the Japanese electoral study, which corresponds to the falling slope in Figure 3 indicating that the probability of giving a correct answer decreases with increasing political knowledge. However, the large standard error of this estimate serves as a warning against such a conclusion. Nevertheless, the low absolute values of several of the estimated item discriminances lead to serious doubts about either the reliability or validity of the corresponding knowledge questions.

3 How Format and Topic Affect Item Performance

The survey questions used in the components of the CSES can be distinguished in terms of their *format*, the the *topic* of the question, and the type of the *required answer*. In terms of format there are *open* and *closed* questions: If respondents are asked to pick one of several alternative, pre-formulated answers, one speaks of a closed format question, otherwise of an open question format. While this distinction is well-known in the survey research literature, the other two dimensions along which the knowledge questions in the CSES surveys can be distinguished are less common. Nevertheless they are potentially relevant for the performance of the items as indicators of political knowledge. How these knowledge items are classified according to format, topic and type of required answer is best explained by some examples.

The American CSES component, the 2004 American National Election Study, uses the following knowledge questions:

1. "Dick Cheney. What job or political office does he NOW hold?" (Correct answer: "Vice-President of the U.S.")
2. "Tony Blair. What job or political office does he NOW hold?" (Correct answer: "Prime Minister of England/Great Britain.")
3. "William Rehnquist. What job or political office does he NOW hold?" (Correct answer: "Chief Justice of the Supreme Court.")

No pre-formulated alternatives are given and thus all three of them are *open*-format questions. In terms of topic they are all questions about *public figures* and they all ask for the name of a *job or office* of the respective public figures.

The French 2002 election study uses the knowledge questions (in the translation given by the CSES general codebook):

1. "Laurent Fabius is a member of the Socialist Party." (Correct answer: "true")
2. "The deputies are elected by proportional representation." (Correct answer: "false")
3. "Michelle Alliot Marie is the president of RPR." (Correct answer: "true")

The questions differ in terms of their topic: the first and the third questions are about public figures, whereas the second question concerns a *political institution*. All three questions require the respondents to assess the *truth-value of a statement*. As a consequence, the format of the questions is *closed*, since only two valid answers are possible. In the German 2002 telephone survey that is used as a component of Module 2 of CSES, the questions employed are:

1. "Who is the current minister of the Interior?"[6] (Correct answer: "Otto Schily")[7]
2. "How many federal states does Germany have after Reunification?"[8] (Correct answer: "16")
3. "How many countries are currently members of the European Union?"[9] (Correct answer: "15")[10]

The first answer is about the *name* of a *public official*, asked in open format. The other two questions concern an *institutional fact* and a fact about *foreign affairs*, respectively, and require the respondents to give a correct *number*. Since none of these questions are administered with pre-formulated choices of answers, they are all in the *open* format.

In the Japanese National Election Study of 2004, the following questions are used (in the translation given by the CSES general codebook):

1. "Which of the following is a requirement to make an amendment to the constitution?"
   (a) "A majority of more than two-thirds of all of the members in both Houses."
   (b) "A majority of more than half of all of the members in both Houses."
   (c) "A majority of more than two-third of all of the members present in both Houses."
   (d) "A majority of more than half of all of the members present in both Houses."
   (*a* is the correct answer.)
2. "Which of the following is one of the requirements to become a Prime Minister in Japan?"
   (a) "Must be in the House of Councilors."
   (b) "Must be part of Congress."
   (c) "Does not necessarily have to be a part of Congress."
   (*b* is the correct answer.)
3. "Out of the following, which is the name given for the election system for the House of Councilors?"
   (a) "Multiple seat constituency system."
   (b) "System of proportional representation as a major part of the system which is combined with single-seat constituencies."

[6] This is the translated wording of the survey question as reported in the German codebook (Weßels and Schmitt 2003). The wording reported in the general CSES codebook differs: The wording reported there is "Identify the Minister of Foreign Affairs."

[7] The CSES code book also gives no information of the correct answers, which are included here, translated from the German codebook.

[8] The wording reported in the general CSES code book is "How many states are there in Germany?"

[9] The wording reported in the general CSES code book is "How many member states are there in the EU?"

[10] The number of EU member states increased to 25 not earlier than 2004 and to 27 after 2007.

Table 2: Distribution of topic, format and types of required answer of knowledge items in CSES Module 2

(a) Topic of question

	Percentage	Count
Institutional	31.5	34
Public official	29.6	32
Public figure	10.2	11
Foreign	8.3	9
Other	20.4	22

(b) Item format

	Percentage	Count
Closed	23.1	25
Open	76.9	83

(c) Type of required answer

	Percentage	Count
Name	50.0	54
Numerical	24.1	26
Other	25.9	28

(c) "System that combines single-seat districts and proportional-seat representation."

(*c* is the correct answer.)

This is a closed-format question asking about an institutional fact, which requires respondents to choose a *correct statement*.

These examples give only a first impression of the variation in topic, answer type and format to be found among the knowledge questions in the components of CSES. Of course, a comprehensive discussion of these aspects of the knowledge questions in the CSES cannot be presented within the limits of this paper. So it has to suffice to report about the classification of knowledge questions according to topic, format, and type of required answer. Since the total number of items is only 108, the categories of question topic and type of required answer are collapsed into five and three categories, respectively, in order to allow for a statistical analysis. The reduced set of topic categories consists of "Institutional facts", "Public official", "Public figure" (other than a public official or not described in terms of her/his public office), and "Foreign affair facts" (including facts about international politics), and "Other". The reduced set of type categories consists of "Name" (of an institution, individual or party), "Numerical" (number or numerical range), and "Other". Table 2 shows how the categories of topic, format, and type of required answer are distributed among the items. Table 5 in the appendix shows how the items are assigned to these categories.

Table 3 shows how question topic, question format, and type of required answer affect the performance of knowledge items in terms of difficulty and discriminance. The table reports estimates of a second-stage linear model of item difficulty and item discriminance, respectively, on topic, format, and answer type. The GLS procedure employed (Long Jusko and Phillips Shively 2005) is designed in such a way that it controlls for the uncertainty about the difficulty and discriminance parameters that are obtained from the logistic latent-trait model described in the previous section and for the correlations among item difficulties

Table 3: Impact of question content and answer format on item constants and discriminances of knowledge questions in CSES2 survey components; generalized least squares estimates of a linear model.

	Difficulty	Discriminance
Intercept	−0.08	1.07***
	(0.18)	(0.19)
Topic: Institutional/Other	0.10	−0.06
	(0.19)	(0.21)
Topic: Public official/Other	0.29*	0.44*
	(0.17)	(0.20)
Topic: Public figure/Other	0.01	−0.10
	(0.23)	(0.26)
Topic: Foreign/Other	0.20	0.50*
	(0.24)	(0.28)
Format: Open/Closed	−0.49*	1.10***
	(0.24)	(0.28)
Type of answer: Name/Other	−0.13	−0.57*
	(0.22)	(0.27)
Type of answer: Numerical/Other	0.86***	−0.76**
	(0.26)	(0.30)
R-squared	0.27	0.45
Log-likelihood	−80.2	−33.5
N	108	108

Note: The weighting matrix is constructed from the inverses of the asymptotic covariances of the estimates of item constants and discriminances. Standard errors in parentheses, significance levels: $***\equiv p < 0.001$, $**\equiv p < 0.01$, $*\equiv p < 0.05$.
Source: Comparative Study of Electoral Systems (CSES) – Module 2 (2001-2006), Full Release – June 27, 2007.

and item discriminances that are used in the same country component of the CSES. The estimation is based on a generalized least squares procedure described in the appendix.

As the estimates reported in Table 3 make clear, the topic of a knowledge question, the open or closed format, and the type of answer it requires impinge both on its difficulty and its discriminance, that is, how well it represents an underlying dimension of political knowledge. Obviously, it makes a considerable difference for both item difficulty and item discriminance if a knowledge question concerns a public official. The estimates indicate that such items are more difficult and discriminate better for political knowledge then do items with other topics. Only items that concern foreign policy topics have a better performance in terms of discriminance.

Item format, however, has the strongest impact on item discriminance. The corresponding coefficient in the equation for item discriminance is slightly larger than the intercept.

Since dummy coding used, this means that, other things being equal, the discriminance of a knowledge question in open format will on average have a discriminance twice as good as a knowledge question in closed format. Also open questions seem to be, on average, slightly less difficult than closed questions, as indicated by the coefficient of item format in the equation for item difficulty.

That the closed-format questions, presenting respondents with a fixed number of given alternatives, have a lower discriminance than open-format questions, comes hardly as surprise. Indeed, the possibility of guessing a correct answer by randomly choosing one of the given alternatives is a concern for educational testing by multiple choice tasks: Test subjects that do not have the appropriate level of knowledge that the task should require nevertheless accidentally choose the correct alternative, with the consequence that their level of cognitive ability is over-estimated. The incidence of such accidentally "correct" answers in an educational test will of course render a test item less effective. That is, multiple choice tasks can be expected to have a lower discriminance for ability than tasks that do not involve given alternatives. But a side-effect of guessing correct answers will be that test items in closed or multiple choice format will be easier than items in open format with the same topic. The results in Table 3, however, run contrary to this reasoning: Closed-format knowledge questions turn out to be more difficult than open-format questions. One explanation for this may be that the stakes for respondents in an election study survey are different than in an educational test. Individuals subjected to educational testing have more to loose from wrong answers than respondents in a survey. That is, one can assume that they will take all cognitive effort they can muster in trying to solve the tasks involved in the test correctly. In a survey context, respondents will not loose much by giving wrong answers (Prior and Lupia 2008). As a consequence, they may be tempted to avoid cognitive effort by randomly selecting one of the alternatives given in a close-format question, even if they could have found a correct answer if the question would have been presented to them in an open format. As a result, a respondent with a given level of political knowledge will, other things being equal, be more likely to give an incorrect answer to a closed-format question than to an open-format question, leading to a higher estimated difficulty parameter for this item.

Not only the format of a knowledge question has an impact on its difficulty and discriminance as an test item of political knowledge, but also the type of answer it requires. As the estimates of the coefficients in the discriminance equation make clear, questions presenting respondents with a question that requires them to give the name of an individual, institution, or party have a lower discriminance an so have questions requiring respondents to give an answer that involves a number or numeric range. An explanation for this may be that the

ability to answer these questions correctly depends not only on general political knowledge, but is also influence by other factors. For example, if the correct answer is the leader of a specific party, this leader may cognitively available to the partisans of this party with higher probability than to those of other parties. Also, partisans may be tempted to attribute certain properties to the party or party leaders they favor.

Questions that require answers that involve a number or range of numbers however, are not only less discriminant for political knowledge than other questions, but also more difficult. If one follows Kuklinski *et al.* (2000), one may consider that numerical estimates are especially affected by a phenomenon that one may describe as "motivated" misinformation. Kuklinski *et al.* suggest that many American citizens over-estimate the amount of federal budget spent on welfare *as a consequence* of a dislike of welfare programs. Similarly motivated estimation mistakes may also be involved in the patterns of answers to political knowledge questions that require stating numbers or numerical ranges. However, after a closer look at these questions used in the CSES Module 2 surveys, this interpretation seems less likely. Rather, one may speculate that questions not only tap respondents' political knowledge, but also their ability of or interest in memorizing and recalling numbers. For example, it is not implausible that even politicos may have difficulty to remember whether the current number of EU-member states is 25, 26, or 27. Even among citizens with a high level of political awareness there may be a substantial proportion who will feel that such numbers not important enough to memorize because they deem them as inessential for the understanding current EU politics.

There are good reasons to attribute the lack of discriminance of knowledge questions asking for numerical values on their lack of validity. They tap another cognitive competence in addition to political awareness. Although less straightforward, the same can be surmised about questions that ask respondents for names of parties, institutions or politicians: The correct answer of such a question requires, beside a certain level of political sophistication, also the ability to remember and recall names. That this requirement poses a challenge of its own is a fact that any instructor will be aware of at least during the first few sessions of his or her course. Consider for example a change in one of the political knowledge questions used in the American National Election Study: Instead of asking respondents about the current job or political office of William Rehnquist, respondents had been asked who the current Chief Justice of the Supreme Court is. In this version of the question respondents would have had to fill in the "non-political" information contained in the statement "William Rehnquist is Chief Justice of the Supreme Court", that is, the Chief Justice's name, whereas in the original version of this question, the non-political information of this sentence is given to

the respondents, and they have to complete the "political" information of this statement, that is, the denomination of William Rehnquist's current office. This example may help making understandable why it is not only crucial what kind of fact about politics is the subject of a political knowledge question but also, what aspect of this fact it asks for. Both variants, the actually used one and the hypothetical one in this example, tap the same fact. Yet by asking about a different aspects of this fact, the hypothetical variant of the ANES question gives the question a different topic, with likely consequences of the discriminance of this question for political awareness.

4 Summary and Conclusions

The present paper raises serious doubts about the equivalence of knowledge questions employed in the election studies that contribute to the CSES. The number of correct responses two the knowledge question batteries varies considerably across samples, even if controlled for education, probably the most powerful predictor of political knowledge. That is, if an individual answers e.g. two political knowledge questions in the Swiss national electoral study this hardly indicates the same level of political knowledge as that of an individual who answers two political knowledge correctly in the Irish national election study. Given the multitude of potential topics of knowledge questions—specific political issues, party leaders, denominations of political institutions, structures of political institutions, electoral thresholds—it seems difficult to gauge the rates of correct responses in advance in such a way that they are identical across countries. But the results of this paper show that there are *systematic* relations between the topic area that is tapped by a knowledge question, by its format (open or closed), and by the type of answer it requires. Questions about public officials relatively often lead to incorrect answers as do questions in closed format and questions requiring respondents to give a number or numerical range.

The IRT-based analyses presented in this paper also indicate that some of the item batteries do not scale very well, that is, several of the items in these batteries have disappointing and even poor performance as discriminators of different levels of political knowledge. That is, the items lack in reliability or validity. However, this does not imply that a concept such as political knowledge does not "travel". Also, the variation in item performance cannot solely attributed to a multi-dimensionality of political knowledge.

It is, of course, quite plausible that political knowledge cannot be strictly uni-dimensional, if Converse's notion of multiple "issue publics" is valid (Converse 1964). In that case, one will be presented with a dilemma: Either knowledge questions will tap more than one dimension

or one will measure only the knowledge of one specific issue public. Yet this cannot be the reason of the sub-optimal performance of the knowledge items in the CSES components, because issue positions of parties or other political actors are rarely used in the CSES knowledge questions. Indeed, the knowledge questions are much more diverse in terms of topics than that. But this diversity alone also cannot be the sole reason for the problems of equivalence and reliability or validity. Instead, there is evidence of a systematic relation between between topic areas and item discriminance: Knowledge questions seem to have a higher discriminance if they concern public officials or foreign affairs.

But it is not only the topic of knowledge questions that accounts for variation among the difficulty and discriminance of knowledge questions. The format of the items and the type of answer the questions require are consequential difficulty and discriminance. A closed format obviously leads respondents into temptation to answer questions by guessing, leading to losses in item discriminance. The tendency to guess, however, has a different effect than in educational testing: Some respondents seem to use guessing to avoid cognitive effort even if they are able to answer questions correctly. As a consequence, the difficulty of closed-format knowledge questions seems to increase. Questions that require respondents to give a name (of an institution, party, or individual) seem discriminate less than other knowledge questions, as do questions requiring respondents to give a numeric response. Such questions also seem to increase the difficulty of the question, that is, increase the probability of an incorrect answer.

In sum, the problem of equivalence with regards to political knowledge questions employed in the CSES is not a "typical" one, in which the literal translation of survey questions nevertheless would not guarantee that they tap the same topics or have the same meaning (Anderson 1967). Rather, the differences in rates of correct responses and item scalability is a consequence of methodological issues. In order to achieve equivalent measures of political knowledge with a limited set of items, CSES participants need to coordinate with regards to the general topical area of the questions, should always use items in open format and should by all means avoid questions that require numbers or statements involving numbers as an answer.

The knowledge questions used in the Amerian National Election Studies are probably a good standard. On the one hand, the three items used in these studies differ in terms of difficulty as one would expect from a good cognitive ability test. On the other hand, the ANES item set is the only one in which all three items achieve a high coefficient of discriminance. The ANES knowledge question ask respondents about the job or office of three different public officials or public figures. These question seem well suited for gauging

discriminance and difficulty: It is easy to construct variants of such questions by exchanging the persons mentioned in these questions. In pilot-studies, several of such variants can be tried out and those variants that achieve the "optimal" rates of correct answers for a three-item battery, that is, 75, 50, and 25 per cent correct answers, respectively, can be selected for the final survey questionnaire. Of course, one may object that the good scalability of the ANES items is an artifact that arises from their similarity in structure and topic. But such an objection would be valid only if these items would tap a very narrow subject matter, say parties' or candidates' positions in a specific policy area. Yet the persons mentioned in the ANES questions are quite heterogeneous: the American Vice President, a foreign head of government, and a Chief Justice of the Supreme Court. Rather than being too specific, they tap simultaneously two aspects of political awareness: To answer the questions correctly, respondents must have followed politics close enough to be able to recognize the mentioned individuals, and they must have sufficient knowledge of political institutions to identify the position or job these individuals have. Of course, since the ANES are the only electoral studies that use knowledge questions of this type, it is impossible to give conclusive systematic evidence for their performance in terms of a comparative analysis as given in the preceding section. But a comparison of these items in terms of their discriminance and difficulty with those used in other electoral studies participating in the CSES is quite suggestive about their performance.

Appendix

Educational Categories in the CSES, Module 2

CSES, Module 2, employs a comprehensive schema for coding levels of educational achievement. The categories of this schema are shown in the left column of Table **??**. In several countries, education was measured using a coarser schema of categories than that of the CSES. In order to allow for educational categories that are more or less comparable, a reduced schema was used for the stratification for education in Figure 2. How this reduced schema relates to the original schema is shown in the right column of Table **??**.

Classification of CSES, Module 2, Knowledge Items

The items used in Module 2 of the Comparative Study of Electoral Systems to measure political knowledge come in a wide variety of question topic, item format, and type of required answer. In order to examine the effects of these properties, the items are grouped

Table 4: The CSES coding schema for education and the reduced set of categories used in this paper

Original CSES	Reduced category set
1 None 2 Incomplete primary	1 Primary incomplete/none
3 Primary completed 4 Incomplete secondary	2 Primary complete
5 Secondary completed 6 Post-secondary trade/vocational school	3 Secondary complete
7 University undergraduate degree incomplete 8 University undergraduate degree completed	4 College/University enrollment
9 [Other]	* [Missing]

into several categories, such that these categories each contain enough cases such that reliable statistical inferences are possible and that they tap the most important aspects of these items. In terms of question topic the items are grouped into the categories *institutional facts*, *public official*, *public figure*, *fact of politics in foreign countries or international relations*, and *Other*. In terms of item format, there are only two categories, *closed* format items that present respondents with any pre-determined alternatives and *open* format items that do not. Examples of the grouping of items in these categories are already given in the main part of the paper. A detailed report of the grouping of the individual knowledge items in the CSES, Module 2 are given in Table 5.

Estimation of IRT Model Parameters

The marginal maximum likelihood procedure used to get the estimates reported in Table 3 treats the individuals' levels of political knowledge as unobserved data with a standard normal distribution—the usual assumption used in IRT modelling (McDonald 1999; van der Linden and Hambleton 1997).

If i is an index denoting an individual and j an index corresponding to an item and y_{ij} is the observed response (one if individual i answers the knowledge question j correctly and zero otherwise), then the marginal log likelihood has the form

$$\ell = \sum_i \log \int_{-\infty}^{\infty} \prod_j \frac{\exp(\beta_j(u_i - \alpha_j))^{y_{ij}}}{1 + \exp(\beta_j(u_i - \alpha_j))} f_N(u_i)\,\mathrm{d}\,u_i \tag{2}$$

Table 5: Classification of CSES2 knowledge items in terms of question content and answer format.

Study component	Item 1			Item 2			Item 3		
	Topic	Format	Type	Topic	Format	Type	Topic	Format	Type
Albania (2005)	Public official	Open	Name	Institutional	Closed	Other	Institutional	Open	Numerical
Australia (2004)	Public official	Open	Name	Other	Open	Name	Other	Closed	Other
Belgium (2003)	Other	Open	Other	Institutional	Closed	Numerical	Institutional	Closed	Numerical
Brazil (2002)	Public official	Open	Name	Public official	Open	Name	Public official	Open	Name
Canada (2004)	Public figure	Open	Name	Other	Open	Name	Other	Open	Name
Chile (2005)	Public official	Open	Name	Public official	Open	Name	Other	Open	Numerical
Finland (2003)	Public figure	Open	Name	Other	Closed	Other	Foreign	Closed	Other
France (2002)	Public figure	Closed	Other	Institutional	Closed	Other	Public figure	Closed	Other
Germany (2002 phone)	Public official	Open	Name	Institutional	Open	Numerical	Foreign	Open	Numerical
Great Britain (2005)	Institutional	Closed	Other	Institutional	Closed	Other	Institutional	Closed	Other
Hong Kong (2004)	Institutional	Closed	Other	Institutional	Closed	Other	Institutional	Closed	Other
Hungary (2002)	Public figure	Closed	Other	Institutional	Open	Numerical	Public official	Open	Name
Ireland (2002)	Public figure	Open	Name	Public figure	Open	Name	Public official	Open	Name
Israel (2003)	Other	Open	Name	Public official	Open	Name	Institutional	Open	Numerical
Italy (2006)	Public official	Open	Name	Public official	Open	Name	Institutional	Open	Numerical
Japan (2004)	Institutional	Closed	Other	Institutional	Closed	Other	Institutional	Closed	Other
Kyrgyzstan (2005)	Public official	Open	Name	Public official	Open	Name	Public official	Open	Name
Mexico (2003)	Institutional	Open	Name	Institutional	Open	Numerical	Public official	Open	Name
Netherlands (2002)	Public official	Open	Name	Public official	Open	Name	Public official	Open	Name
New Zealand (2002)	Institutional	Closed	Other	Institutional	Closed	Other	Other	Open	Name
Norway (2001)	Public official	Open	Name	Other	Open	Name	Public official	Open	Name
Peru (2006)	Public official	Open	Name	Other	Open	Numerical	Other	Open	Name
Philippines (2004)	Institutional	Open	Numerical	Institutional	Closed	Name	Public figure	Open	Other
Poland (2001)	Foreign	Open	Other	Public official	Open	Name	Foreign	Open	Name
Portugal (2002)	Other	Open	Name	Foreign	Open	Numerical	Public figure	Open	Name
Portugal (2005)	Other	Open	Name	Other	Open	Name	Foreign	Open	Numerical
Republic of Korea (2004)	Institutional	Open	Numerical	Institutional	Open	Numerical	Public official	Open	Name
Romania (2004)	Institutional	Open	Numerical	Public official	Open	Name	Institutional	Open	Numerical
Russia (2004)	Other	Open	Name	Public official	Open	Name	Public official	Open	Name
Slovenia (2004)	Public official	Open	Name	Foreign	Open	Numerical	Institutional	Open	Numerical
Spain (2004)	Other	Open	Name	Other	Open	Numerical	Foreign	Open	Numerical
Sweden (2002)	Other	Closed	Other	Other	Closed	Other	Other	Closed	Other
Switzerland (2003)	Public official	Open	Name	Institutional	Open	Numerical	Institutional	Open	Numerical
Taiwan (2001)	Public official	Open	Name	Other	Open	Name	Institutional	Open	Numerical
Taiwan (2004)	Foreign	Open	Name	Institutional	Open	Name	Institutional	Open	Numerical
United States (2004)	Public figure	Open	Other	Public figure	Open	Other	Public official	Open	Other

Note: No knowledge questions were asked in the Bulgaria (2001), Germany (2002) mailback, Denmark (2001), and Iceland (2003) survey components.

where $f_N(u_i)$ is the value of the standard normal density distribution. Since the integral involved in the marginal log-likelihood is a univariate Gaussian integral, it can be effectively approximated by Gauss-Hermite quadrature. The procedure used in this paper employs 21 quadrature points (Rizopoulos 2008).

Second-Stage Regression of IRT Model Estimates

In the third section of this paper, second-stage regression models used in this paper to examine the effects of several aspects of the knowledge questions on their difficulty and their discriminative power to distinguish between levels of political knowledge. The dependent variables in these second-stage regression models are, respectively, the estimates of the difficulty and discriminance parameters obtained in the second section. The second-stage regression model is constructed in such a way that it accounts for the uncertainty associated with these estimates.

Let $\boldsymbol{b}_j$ denote the vector of "true" parameter values for country j of the three discriminance or difficulty parameters. The second stage regression model can then be written as

$$\boldsymbol{b}_j = \boldsymbol{X}_j\boldsymbol{\gamma} + \boldsymbol{u}_j \tag{3}$$

where $\boldsymbol{X}_j$ is a design matrix composed of the values of the dummy variables for the topic and format of the knowledge questions and the type of answer they require, $\boldsymbol{\gamma}$ is a coefficient vector, and $\boldsymbol{u}_j$ is a vector of residuals with zero mean and variance σ^2 (Long Jusko and Phillips Shively 2005).

Provided that the estimation procedure of the IRT model used in this paper leads to asymptotically normal distributed estimates $\hat{\boldsymbol{b}}_j$, they have zero expectation and asymptotic covariance matrix $\boldsymbol{W}_j$. The second-stage log-likelihood then is

$$L = -\frac{mn}{2}\log(2\pi) - \sum_j \frac{1}{2}\log|\boldsymbol{V}_j| - \sum_j \frac{1}{2}(\hat{\boldsymbol{b}}_j - \boldsymbol{X}_j\boldsymbol{\gamma})'\boldsymbol{V}_j^{-1}(\hat{\boldsymbol{b}}_j - \boldsymbol{X}_j\boldsymbol{\gamma}) \tag{4}$$

with $\boldsymbol{V}_j = \hat{\boldsymbol{W}}_j + \boldsymbol{I}_j\sigma^2$, where $\hat{\boldsymbol{W}}_j$ is the estimated asymptotic covariance matrix of $\hat{\boldsymbol{b}}_j$, $\boldsymbol{I}_j$ is an identity matrix of appropriate size, m is the number of items, and n is number of countries.

Evidently, for given σ^2, estimates for the second-stage coefficients $\boldsymbol{\gamma}$ can be found by minimizing the sum of squares

$$S = \sum_j (\hat{\boldsymbol{b}}_j - \boldsymbol{X}_j\boldsymbol{\gamma})'\boldsymbol{V}_j^{-1}(\hat{\boldsymbol{b}}_j - \boldsymbol{X}_j\boldsymbol{\gamma}) \tag{5}$$

or, equivalently, by solving the GLS equation

$$\sum_j \boldsymbol{X}_j' \boldsymbol{V}_j^{-1} \boldsymbol{X}_j \boldsymbol{\gamma} = \sum_j \boldsymbol{X}_j' \boldsymbol{V}_j^{-1} \hat{\boldsymbol{b}}_j \tag{6}$$

for $\boldsymbol{\gamma}$.

The appropriate σ^2 can be found by solving the ANOVA equation (Jiang 2007: 27)

$$\sum_j \hat{\boldsymbol{b}}_j' \hat{\boldsymbol{W}}_j^{-1} \boldsymbol{P}^{\perp}_{\boldsymbol{X}_j, \boldsymbol{W}_j^{-1}} \hat{\boldsymbol{b}}_j = \sigma^2 \sum_j \operatorname{tr}\left(\hat{\boldsymbol{W}}_j^{-1} \boldsymbol{P}^{\perp}_{\boldsymbol{X}_j, \boldsymbol{W}_j^{-1}} \right) \tag{7}$$

where $\boldsymbol{P}^{\perp}_{\boldsymbol{X}, \boldsymbol{W}^{-1}}$ is the residual projection matrix

$$\boldsymbol{P}^{\perp}_{\boldsymbol{X}, \boldsymbol{W}^{-1}} = \boldsymbol{I} - \boldsymbol{X} \left(\boldsymbol{X}' \boldsymbol{W}^{-1} \boldsymbol{X} \right)^{-1} \boldsymbol{X}' \boldsymbol{W}^{-1} \tag{8}$$

The estimation of the second-stage regression coefficients thus proceeds in two steps. In the first step, the second-stage residual variance σ^2 is estimated based on equation (7), and in the second step, the regression coefficient parameter estimates are computed by solving the GLS equation (6). An implementation of this GLS second-stage regression estimation procedure is written in *R* and is available from the author of this paper on request. It will be made publicably available once the paper is published.

References

Anderson, R. Bruce W. 1967. "On the Comparability of Meaningful Stimuli in Cross-Cultural Research." *Sociometry* 30(2):124–136.

Bartholomew, David J., Fiona Steele, Irini Moustaki, and Jane I. Galbraith. 2008. *Analysis of Multivariate Social Science Data*. Boca Raton: Chapman & Hall/CRC Press, second edition.

Berelson, Bernard R., Paul F. Lazarsfeld, and William N. McPhee. 1954. *Voting: A Study of Opinion Formation in a Presidential Campaign*. Chicago: University of Chicago Press.

Berggren, Heidi M. 2001. "Institutional Context and Reduction of the Resource Bias in Political Sophistication." *Political Research Quarterly* 54(3):531–552.

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1960. *The American Voter*. New York: Wiley.

Converse, Philip E. 1964. "The Nature of Belief Systems in Mass Publics." In *Ideology and Discontent*, ed. David E. Apter. New York: The Free Press, pp. 206–261.

Converse, Philip E. 1990. "Popular Representation and the Distribution of Information." In *Information and Democratic Processes*, eds. John A. Ferejohn and James H. Kuklinski. Urbana and Chicago: University of Illinois Press, pp. 369–388.

Converse, Philip E. 2000. "Assessing the Capacity of Mass Electorates." *Annual Review of Political Science* 3:331–353.

CSES. 2007. "The Comparative Study of Electoral Systems Module 2 Full Release. June 27, 2007 version: Code book."
URL `http://www.cses.org`

Delli Carpini, Michael X. and Scott Keeter. 1996. *What Americans Know about Politics and Why it Matters*. New Haven and London: Yale University Press.

Gibson, James L. and Gregory A. Caldeira. 2008. "Knowing the Supreme Court? A Reconsideration of Public Ignorance of the High Court."
URL `http://polisci.wustl.edu/sub_page.php?s=3\&m=0\&d=7`

Gordon, Stacy B. and Gary M. Segura. 1997. "Cross-national Variation in the Political Sophistication of Individuals: Capability or Choice?" *Journal Of Politics* 59(1):126–147.

Grönlund, Kimmo and Henry Milner. 2006. "The Determinants of Political Knowledge in Comparative Perspective." *Scandinavian Political Studies* 29(4):386–406.

Huckfeldt, Robert and John Sprague. 1995. *Citizens, Politics, and Social Communication: Information and Influence in an Election Campaign*. Cambridge: Cambridge University Press.

Jerit, Jennifer, Jason Barabas, and Toby Bolsen. 2006. "Citizens, Knowledge, and the Information Environment." *American Journal of Political Science* 50(2):266–282.

Jiang, Jiming. 2007. *Linear and Generalized Linear Mixed Models and Their Applications*. New York: Springer.

King, Gary, Christopher J. L. Murray, Joshua A. Salomon, and Ajay Tandon. 2004. "Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research." *American Political Science Review* 98(1):191–2007.

Krosnik, Jon A., Arthur Lupia, Mathew DeBell, and Darrell Donakowski. 2008. "Problems with ANES Questions Measuring Political Knowledge." Note of the ANES Principal Investigators and Operational Directors.

Kuklinski, James H. and Paul J. Quirk. 2001. "Conceptual Foundations of Citizen Competence." *Political Behavior* 23(3):285–311.

Kuklinski, James H., Paul J. Quirk, Jennifer Jerit, and Robert F. Rich. 2001. "The Political Environment and Citizen Competence." *American Journal of Political Science* 45(2):410–424.

Kuklinski, James H., Paul J. Quirk, Jennifer Jerit, David Schwieder, and Robert F. Rich. 2000. "Misinformation and the Currency of Democratic Citizenship." *The Journal of Politics* 62(3):790–816.

Lau, Richard R. and David P. Redlawsk. 2008. "Advantages and Disadvantages of Cognitive Heuristics in Political Decision Making." *American Journal of Political Science* 45(4):951–971.

Long Jusko, Karen and W. Phillips Shively. 2005. "Applying a Two-Step Strategy to the Analysis of Cross-National Public Opinion Data." *Political Analysis* 13(4):327–344.

Lupia, Arthur and Mathew D. McCubbins. 1998. *The Democratic Dilemma: Can Citizens Learn What They Need To Know?* Cambridge: Cambridge University Press.

Luskin, Robert C. 1987. "Measuring Political Sophistication." *American Journal of Political Science* 31(4):856–899.

McDonald, Roderick P. 1999. *Test Theory: A Unified Treatment*. Mahwah, New Jersey and London: Lawrence Erlbaum.

Mondak, Jeffery J. 2000. "Reconsidering the Measurement of Political Knowledge." *Political Analysis* 8(1):57–82.

Mondak, Jeffery J. 2001. "Developing Valid Knowledge Scales." *American Journal of Political Science* 45(1):224–238.

Neuman, Russel W. 1986. *The Paradox of Mass Politics: Knowledge and Opinion in the American Electorate*. Cambridge, Mass.: Harvard University Press.

Prior, Markus. 2007. *Post-broadcast Democracy: How Media Choice Increases Inequality in Political Involvement and Polarizes Elections*. Cambridge: Cambridge University Press.

Prior, Markus and Arthur Lupia. 2008. "Money, Time, and Political Knowledge: Distinguishing Quick Recall and Political Learning Skills." *American Political Science Review* 52(1):169–183.

Przeworski, Adam and Henry Teune. 1966. "Equivalence in Cross-National Research." *Public Opinion Quarterly* 30:551–568.

R Development Core Team. 2008. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
URL `http://www.R-project.org`

Rizopoulos, Dimitris. 2008. "ltm: Latent Trait Models under IRT (Ver. 0.8-4)."
URL `http://cran.r-project.org/web/packages/ltm/`

Sturgis, Patrick, Nick Allum, and Patten Smith. 2008. "An Experiment on the Measurement of Political Knowledge in Surveys." *Public Opinion Quarterly* 85(1):90–102.

van der Linden, Wim J. and Ronald K. Hambleton, eds. 1997. *Handbook of Modern Item Response Theory*. New York and Berlin: Springer.

Verba, Sidney. 2003. "Would the Dream of Political Equality Turn out to Be a Nightmare?" *Perspectives on Politics* 1(4):663–679.

Weßels, Bernhard and Herman Schmitt. 2003. "Deutsche Nationale Wahlstudie - Nachwahlstudie 2002 Deutsche CSES-Studie (CSES-Modul II)."
URL `http://www.wzb.eu/~wessels/DSL-download.en.htm`