Changes in Knowledge about and Perception of Civics and Citizenship over a Ten-Year Period: Comparing CIVED 1999 and ICCS 2009

Wolfram Schulz
Australian Council for Educational Research
Melbourne/Australia
schulz@acer.edu.au

Paper presented to the symposium on "Evaluating the Citizenship Competences of Young Adolescents in Europe: Methods, Patterns and Trends"
at the European Conference on Educational Research (ECER)
in Cádiz, 18-21 September 2012.

Introduction

A larger number of countries involved in the IEA International Civic and Citizenship Education Study (ICCS 2009) had also participated in the previous IEA CIVED study undertaken in 1999. The civic knowledge test included a subset of CIVED items which were used to derive comparable measures of civic content knowledge, one of the subscales reported in CIVED 1999. Some of the ICCS 2009 questionnaire material also included items that were similar or even identical in wording to those used in CIVED but for which some general format changes were applied. This paper discusses the potential for comparing results from both surveys for test and questionnaire items and discusses the caveats researchers have to keep in mind when doing so. The paper will illustrate this with selected results from European countries from lower secondary students between 1999 and 2009.

The first part of this paper describes the equating analyses and procedures undertaken for common cognitive items as well as European results across both surveys overall and for selected sub-groups (e.g. gender, books at home). The second part will discuss the possibilities of comparing questionnaire results between the two surveys in view of the format changes that were applied to the material. The conclusion will summarise the results and discuss the implications of design, format and content changes for comparing data from CIVED 1999 and ICCS 2009.

The International Civic and Citizenship Education Study (ICCS)

ICCS 2009 was the third international IEA study designed to measure context and outcomes of civic and citizenship education and it was explicitly linked through common questions to the IEA Civic Education Study (CIVED) which was undertaken in 1999 and 2000 surveying 14-year-old lower secondary students and upper secondary students between the ages of 16 and 18 (Torney-Purta, Lehmann, Oswald and Schulz, 2001; Amadeo et. al., 2004; Schulz and Sibberns, 2004). ICCS 2009 surveyed 13-to-14-year old students in 38 countries in the years 2008 and 2009 and reported on students' civic knowledge, engagement and perceptions as well as on the context for civic and citizenship education (Schulz, Ainley, Fraillon, Kerr & Losito, 2010a & 2010b). Outcome data were obtained from representative samples of students in their eighth year of schooling and context data from the students, their schools and teachers. In addition, an on-line survey carried out through national centres informed on the context of civic and citizenship education at the national level.[1]

In addition to the international test and questionnaires, regional instruments were administered in Asia, Europe and Latin America. These instruments consisted of short knowledge tests and questionnaire material designed to capture region-specific knowledge and perceptions. The results of the regional surveys were published in a series of regional reports (Kerr, Sturman, Schulz & Burge, 2010; Schulz, Ainley, Friedman & Lietz, 2011; Fraillon, Schulz & Ainley, 2012) and were also included in numerous reports and publications within countries.

[1] Further information about ICCS can be found at its website http://iccs.acer.edu.au/.

It is well known that there is substantial diversity in the field of civic and citizenship education within and across countries. Consequently, maximising the involvement of researchers from participating countries in this international comparative study was of particular importance for the success of this study in the process of developing an assessment framework and instruments. As for CIVED 1999 and other IEA studies, the international study centre for ICCS 2009 sought input from national research centres throughout the study and strategies were developed to maximise country contributions from early piloting activities until the selection of final main survey instruments in June 2009.

The students surveyed for ICCS were students enrolled in the grade that represents eight years of schooling, counting from the first year of ISCED Level 1, provided the mean age at the time of testing is at least 13.5 years. According to this definition, for most countries the target grade was the eighth grade, or its national equivalent.

The aim of the survey was to gather data on (a) student knowledge, conceptual understanding and competencies in civic and citizenship education, (b) student background characteristics and participation in active citizenship, and (c) student perceptions of aspects of civics and citizenship. Instruments used in ICCS included an on-line national context survey completed by national centres, a student test, a student questionnaire, a teacher questionnaire and a school questionnaire.

The ICCS assessment framework (Schulz, Fraillon, Ainley, Losito & Kerr, 2008) outlined the aspects that were addressed in the cognitive test and student perceptions questionnaire and provides a mapping of factors that might influence outcome variables and explain their variation. The main data collection took place between October and December 2008 in the educational systems with Southern Hemisphere school calendar year and between February and May 2009 in those with a Northern Hemisphere school calendar year.

The analyses presented in this paper will focus on test and questionnaire data from European countries that participated in both CIVED 1999 and ICCS 2009.

Comparison of civic knowledge over time

Scaling and equating procedures

Seventeen CIVED items were included in the international test to allow the reporting of trends from the previous IEA civic and citizenship education survey in 1999. All CIVED and ICCS test items were scaled using item response theory (IRT) scaling methodology, more specifically the one-parameter (Rasch) model (Rasch, 1960). For link items which were all dichotomous this means that the probability of selecting Category 1 (correct response) instead of 0 (incorrect response) is modelled as

$$P_i(\theta) = \frac{\exp(\theta_n - \delta_i)}{1 + \exp(\theta_n - \delta_i)},$$

where $P_i(\theta)$ is the probability for person n to score 1 on item i, θ_n is the estimated ability of person n, and δ_i is the estimated location of item i on this dimension. For each item, item responses are modelled as a function of the latent trait θ_n. The scaling software package ACER Conquest, Version 2.0 software (Wu, Adams, Wilson, & Haldane, 2007) was used to scale ICCS test and questionnaire data (see further details in Schulz, Ainley & Fraillon, 2011).

Given the change in test design between the two studies, we observed modest positioning effects on item difficulties in the CIVID data for a number of countries. Whereas in CIVED, the assessment consisted of only one booklet where each item appeared in only one position within the test, ICCS used a rotated design, which ensured that students responded to link items in each of the three possible positions at the start, middle, or end of the assessment.

As a consequence, after controlling for ability, the CIVED 1999 students had a higher probability than the ICCS 2009 students of giving correct answers to items that had been administered at the beginning but were less likely to know the answers to test questions administered toward the end of the assessment. However, this effect was not very strong for the pooled international samples and was notable in only a few (mostly non-European) countries.

Given that the framework for the test domain was broader in ICCS 2009 than in CIVID 1999 and given that the majority of the link items represented only one content domain (civic systems and society), it was not surprising that preliminary analysis showed some notable differences in the behaviour of the CIVED link items and the new ICCS items. Effort to estimate comparable ICCS test scores using the newly established scale for the CIVED data was not deemed appropriate.

Despite this consideration, it was decided to test an approach that involved using different equating methods to set test data based on the set of link items (with reduced sample size, given these items appeared in only three out of the seven ICCS booklets) against the CIVED scale metric. However, comparisons of the differences in percentages correct for both surveys and the resulting trend estimates showed several inconsistencies that were probably a consequence of the set of link items including only two items from the sub-dimension "interpretative skills." Consequently, a decision was made to report comparisons only with regard to the "civic content knowledge" subscale established in CIVED, for which 15 link items were available.

In order to review the link item characteristics, we compared the adjusted item difficulty parameters (each standardized to have a mean of 0) first at the international level and then for each national sample. Figure 1 shows the scatter plot between the item parameters from CIVED 1999 and those estimated for the trend sample consisting of 500 randomly selected students from each of the national samples with comparable data. As is apparent in the figure, five items were slightly outside the error bands derived from the respective standard errors of the item parameters from both calibrations. However, the figure also shows that the item parameters were generally highly similar; the correlation between item parameters was 0.96.

Figure 1: Scatter plot for link item parameter estimates from CIVED and for the ICCS trend sample

Some national items from scaling were omitted from scaling including those that had been excluded from the CIVED scaling (see Schulz & Sibberns, 2004), ICCS versions of items with translation errors or deviations, items that reflected printing problems, and items that showed very large differences in relative item difficulty between the two surveys.

For the final scaling, maximum likelihood estimates (MLEs) were computed using the same item parameters as in CIVED, and then transformed them to the same scale metric, which was set to have a mean of 100 and a standard deviation of 20 for the 28

countries that participated in CIVED. Scale scores were transformed to this metric by applying this formula:

$$\theta'_n = 100 + 20\left(\frac{\theta_n - \bar{\theta}}{\sigma_\theta}\right)$$

Here, θ'_n are the student scores in the CIVED metric, θ_n are the original logit scores (maximum likelihood estimates), $\bar{\theta}$ is the CIVED mean of student logit scores (0.95) with equally weighted country subsamples, and σ_θ is its corresponding CIVED standard deviation (1.36). Table 2 shows the item parameters used for scaling as well as the average percentages of correct responses for these items in the 17 countries in 2009 and 1999.

Table 2: Item parameters and average percent correct for link items

Scale scores could be derived only for those students who responded to the link item cluster (included in three out of the seven randomly allocated booklets) and only for those 17 national datasets where the respective student populations were comparable with the ones surveyed in CIVED in 1999. Table 3 records the scale reliabilities (Cronbach's alpha) for this subset of test items as well as the number of items that were used for scaling (after national item exclusions). The median reliability of this set of test items was 0.77, and the reliabilities ranged from 0.69 to 0.82 across the national samples.

Table 3: Test reliabilities for link items (Cronbach's alpha)

Because the transformation equating the ICCS 2009 data with the CIVED 1999 data depended on the change in the degree of difficulty of each of the individual link items, the sample of link items chosen influenced the choice of transformation. This meant that the resulting transformation would have been slightly different if we had chosen an alternative set of link items. Uncertainty in the transformation thus relates to the sampling of the link items, in the same way that uncertainty in values such as country averages is an outcome of the particular sample of students that is used.

The uncertainty resulting from link-item sampling is referred to as linking error, and it is an error that analysts have to take into account when comparing the results arising out of different data collections (see Monseur & Berezner, 2007). As is the situation with the error that is introduced through the process of sampling students, the exact magnitude of this linking error cannot be determined. We can, however, estimate the likely range of magnitudes for this error and take it into account when interpreting results. As with sampling errors, the likely range of magnitude for the errors is represented as a standard error.

Because all link items were dichotomous and not clustered in units, we were able to compute the linking error for ICCS by using the following simple formula:

$$\sigma_{(Linking\_error)} = \sqrt{\frac{\sigma^2}{n}}$$

Here, σ^2 represents the variance of the item parameter differences between 1999 and 2009 (using international calibration samples), and *n* denotes the number of link items used. The linking error for trend reporting from 1999 to 2009 was 0.65 score points in

the final reporting metric (0.044 logits), and we took it into account when estimating the statistical significance of differences.

When testing the difference of a statistic between the two assessments, we computed the standard error of the difference as follows:

$$SE(\mu_{ICCS} - \mu_{CIVED}) = \sqrt{SE_i^{\,2} + SE_j^{\,2} + EqErr^2}$$

Here, μ can be any statistic in units on the ICCS–CIVED link scale (mean, percentile, gender difference, but not percentages) and SE_i and SE_j are the respective standard errors of this statistic from the two surveys. *EqErr* denotes the equating error that reflects the uncertainty in the link between both assessments, which was equal to 0.65 score points on the link scale. Given that the link scale scores were maximum likelihood estimates and not plausible values, we did not have to provide for imputation error when computing the standard errors of the differences in civic content knowledge between 1999 and 2009.

Comparing civic content knowledge in European countries

All participating countries completed the CIVED link items and their item scores contributed to the total ICCS scale scores. Eighteen of the countries that participated in CIVED also participated in ICCS, and 17 of these countries (including two non-European countries) used the same item translations in ICCS as in CIVED in order to permit a comparison of performance across time. As part of the ICCS quality assurance procedures, translations of link items were independently verified and compared against those used in CIVED to ensure a maximum of comparability. Following initial scaling analyses some items were re-verified and in a few cases items were found to have translation errors which were removed from the equating for the respective national samples.

Two countries of 17 countries with comparable target grade data, England and Sweden, tested students at different times of the school year in CIVED and ICCS: England tested its target CIVED grade students (grade 9) at the beginning of the following school year (about half a year later than in ICCS), whereas Sweden undertook its student survey at the beginning of the school year for its target grade (8). Therefore, in England, the students surveyed in CIVED were about half a year older than those surveyed in ICCS, and in Sweden the students who participated in CIVED were about half a year younger than those who participated in ICCS. The results from these countries were reported in a separate section of in the international reports (see Schulz et al., 2010a & 2010b) as it is unknown to which extent differences in the age of the CIVED students and the ICCS students influenced the outcomes. Therefore, for 13 out of 15 European countries that participated in both surveys valid comparisons of performance between 1999 and 2009 could be conducted.

Table 3 shows the comparison between civic content knowledge scores between 1999 and 2009 for European ICCS countries with comparable data. In 1999, the average score on the civic content knowledge scale across the 13 European countries with comparable data was 101 scale points; the average score for the same countries in ICCS 2009 was 98 scale points. This difference translates into a (statistically significant) overall decrease in average performance on the civic content knowledge scale items of four points, or one fifth of a standard deviation.

The average civic content knowledge scale score was statistically significantly higher in ICCS than CIVED in only one country—Slovenia—where the difference was three scale points. In six countries, no statistically significant difference emerged between the 1999 and 2009 scores. The average civic content knowledge scores of six countries decreased statistically significantly between CIVED and ICCS. The largest decrease in performance—11 points—occurred in Bulgaria.

The average age of students across all 13 countries included in the comparison was 14.7 years in CIVED and 14.6 in ICCS; the data in Table 3 show only small differences between student age between CIVED and ICCS data collections.

Table 3: Changes in civic content knowledge between 1999 and 2009 in European ICCS countries

Different background questions were used in CIVED and ICCS 2009 and only some of them are directly comparable. Therefore, comparisons of civic content knowledge results will be compared to three indicator variables:

- Students' sex (male vs. female);
- Students' country of birth (country of test vs. others);
- Students' parents' education (at least one parent completed university vs. others).

Table 4 shows the gender differences (average females minus average males) in civic content knowledge in comparison for the 13 European countries that participated with the same target population and testing windows in both surveys.

Table 4: Gender differences in civic content knowledge in comparison

Generally, for this scale only minor differences between gender groups were recorded. Overall across these 13 European ICCS countries, in 2009 females had slightly but statistically significant higher scale scores than males. However, within countries only three countries had a statistically significant gender difference in favour of male students in CIVED (Czech Republic and Switzerland), there were no significant differences recorded in ICCS 2009. It should be noted that gender differences for the ICCS civic knowledge scale tended to be more consistently in favour of females (see Schulz et al., 2010a & 2010b).

Table 5: Differences in civic content knowledge by country of birth in comparison

Table 5 shows the differences in civic content knowledge scores between students who were born in the country of test and others. In both surveys "native" students tended to have significantly higher scales scores than those born in another country. Across the 13 European countries, the differences were 4.4 points in CIVED 1999 and 4.7 points in ICCS 2009. Whereas in CIVED 1999 statistically significant differences in favour of "native" students were found in eight countries, this was the case in only four countries in ICCS 2009. However, it should be noted that standard errors were considerably high in many countries due to the low percentages of students born in another country which affects tests of significance.

Table 6: Differences in civic content knowledge by parental education in comparison

Table 6 shows differences in scale scores of civic content knowledge between students who reported to have at least one parent with a university degree and other students. The results show a consistently positive and statistically significant effect of

parental education on civic content knowledge. On average, having a parent with a university degree was associated with a difference of 9.5 points in CIVED 1999 and 8.3 points in ICCS 2009. In the Czech Republic, Estonia and Slovenia somewhat smaller effect sizes were found in 2009 in comparison with those from 1999. However, it needs to be taken into account that the question format was different across the two surveys: Whereas in ICCS 2009 categories of parental education were nationally adapted to match the international ISCED classification, directly translated categories were used in CIVED 1999.

For the background variables used in these comparisons it can be concluded that generally similar results were found in both surveys. There is some indication that gender differences slightly changed in favour of females and that differences between native and immigrant students may have changed in some of the European countries. However, these changes were not very large in most cases and interpretations of these results need to take the relatively small number of link items in ICCS 2009 into account.

Comparing questionnaire results from 2009 and 1999

Given that it was conceived as a baseline study for future surveys on civic and citizenship education with links but not as a continuation of CIVED, the development of questionnaires for ICCS 2009 aimed at creating a new set of instruments. Some of the questionnaire material from CIVED was retained but, following discussions with national coordinators and experts, it was also further refined and a different format h was used. Format changes included not to retain a general category for "don't know" and also, in a number of cases, to reverse the categories of Likert-type items. For example, agreement item scales in CIVED started with "strongly disagree" as the lowest category whereas in ICCS they had with "strongly agree" as the first category.

Other differences include modifications of question stem and item content which in some cases were applied to take changes in the general context into account. For example, instead of always referring to "women's rights" when asking students about their attitudes toward this topic, ICCS 2009 focused more on "gender equality" with a series of similar items with only slight changes in wording.

In this section we will present as an example data for four questionnaire items which were designed to about students' attitudes towards equal rights and responsibilities for all ethnic/racial in their countries. In both surveys students were asked to rate their agreement ("strongly agree", "agree", "disagree" and "strongly agree") with the following statements:

- Item 1: All <ethnic/racial groups> should have an equal chance to get a good education in <country of test>;
- Item 2: All <ethnic/racial groups> should have an equal chance to get good jobs in <country of test>;
- Item 3: Schools should teach students to respect <members of all ethnic/racial groups>;
- Item 4: <Members of all ethnic/racial groups> should be encouraged to run in elections for political office.

These four items had identical item wording in both IEA surveys but in ICCS 2009 they were augmented by a fifth item to derive a more reliable scale. Another difference was that in CIVED 1999 the items were included in a larger pool of items asking about different aspects of equal rights and responsibilities (also for women and anti-democratic groups) whereas in ICCS 2009 they were included in one question.

Table 7: Percentages of missing responses for ethnic group items in comparison

Table 7 shows the percentages of missing responses for the items in each survey. Student responses in the "don't know" category were treated as missing responses for the CIVED data. The results illustrate a much higher amount of missing data in 1999 where on average between seven and 19 percent of item responses were missing. The ICCS data, in contract, had only about 2 percent of missing responses. The higher percentages of missing responses in CIVED indicate the effect of offering students a category for "don't know". The even higher missing percentage for the fourth item might be due to the fact that it appeared towards the end of a longer item battery.

Table 8: Percentages of agreement for items 1 and 2 in comparison

Table 8 shows a comparison of item percentages of student agreement ("strongly agree" and "agree") for items 1 and 2 across the two surveys. At both data collections large majorities of students tended to endorsed that all ethnic/racial groups in their countries should have equal chances to education and jobs in their country. Across European countries that participated in both surveys, there was, on average, a slight increase in agreement percentages of about four points. The highest increase in agreement to these items was recorded for Switzerland (German part) where the increase was 15 and 16 percent respectively. In only few countries in 2009 we observed (slightly) lower percentages of agreement than in 1999.

Table 9: Percentages of agreement for items 3 and 4 in comparison

Table 8 shows a comparison of the corresponding item percentages of student agreement ("strongly agree" and "agree") for items 3 and 4. As with the other two items, at both data collections large majorities of students recorded their endorsement that students should teach respect for all ethnic/racial groups in their countries and that members of all ethnic/racial groups should be encouraged to run in elections for political office in their country. The latter item (4) had notably lower levels of endorsement than the former (item 3).

Across European countries that participated in both surveys, there was, on average, again a slight increase in agreement percentages of about four points for item 3. However, item 4 recorded on average a lower level of endorsement in 2009 when compared to 1999. It should, however, be taken into account that in 1999 only four out of five students (on average) gave a response to this item. The highest increase in agreement for item 3 was again recorded for Switzerland (German part) whereas in this country there was also a slight decrease in endorsement of item 4.

When comparing ICCS 2009 questionnaire data with those from CIVED 1999 it is important to do this with caution. The results show some differences between surveys but, in particular for the last item, differences are likely to be affected by the format changes between the two surveys. With regard to other questionnaire item material, it is also important to note that there were quite a few modifications in stem and item content that make it difficult to compare results between the two data collections.

Conclusion and Discussion

This paper discusses the possibilities of comparing data from the CIVED 1999 and the ICCS 2009 survey for participating European countries. Whereas for cognitive test items comparisons were supported by a common set of link items that were translated in exactly the same way, given the nature of establishing ICCS 2009 as a baseline study for future surveys questionnaire items with similar content underwent more considerable modifications, in particular with regarding to the ordering of categories and the omission of a "don't know" category in the ICCS material.

However, even for the comparison of test results it needs to be taken into account that there was a general format change from a single- to a (rotated) multi-booklet design and that there were constraints regarding the availability of items covering all aspects of the civic knowledge content measured in 1999. Consequently, comparisons are restricted to the sub-dimension of civic content knowledge and interpretations of change need to take into account that results may have been affected by item positions in the CIVED assessment. Scale comparisons also depend on a relatively small set of 15 common test items. Therefore it is important to take the equating error into account (as outlined in this paper) when computing the standard errors for the differences in civic content knowledge between the two surveys.

Researchers who wish to make cross-survey comparisons for any questionnaire material are strongly advised to this with utmost caution. The results presented in this paper show the considerable differences in missing percentages for four items measuring attitudes toward equal rights for ethnic/racial groups which did not change in wording. For the fourth item (encouraging members of all ethnic/groups to run in elections for political office) some considerable (mostly negative) changes between the two surveys were found for European participating countries, whereas for other items agreement increased between the two surveys. However, this particular item also substantial proportions of missing responses given its position towards the end of larger item battery. Therefore, it is difficult to make any real inferences from these observations.

Future IEA surveys of civic and citizenship education will include item material for which (a) no format or content modifications will be applied and (b) care will taken to control any possible positioning effects. Therefore it is expected to have cross-sectional comparable data from both test and questionnaires in future ICCS assessments. In the meantime, researchers interested in comparing civic-related IEA survey results over time are advised to undertake this keeping in mind the implications from changes in design, format and content between CIVED 1999 and ICCS 2009 for interpreting results from comparative analyses.

References

Amadeo, J., Torney-Purta, J., Lehmann, R., Husfeldt, V., and Nikolova, R. (2002). *Civic knowledge and engagement: An IEA study of upper secondary students in sixteen countries*. Amsterdam: International Association for the Evaluation of Educational Achievement (IEA).

Fraillon, J., Schulz, W. & Ainley, J. (forthcoming). *ICCS 2009 Asian Report. Civic knowledge and attitudes among lower secondary students in five Asian countries*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Kerr, D., Sturman, L., Schulz, W., & Bethan, B. (2010). *ICCS 2009 European Report. Civic knowledge, attitudes and engagement among lower secondary school students in twenty-four European countries*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Monseur, C. and Berezner, A. (2007). The Computation of Equating Errors in International Surveys in Education. Journal of Applied Measurement, 8(3), 323-335.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Nielsen & Lydiche.

Schulz, W. & Sibberns, H. (2004) (Eds.) *IEA Civic Education Study. Technical Report*. Amsterdam: IEA.

Schulz, W., Ainley, J., Fraillon, J., (Eds.) (2011). *ICCS 2009 Technical Report*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Schulz, W., Ainley, J., Fraillon, J., Kerr, D. & Losito, B. (2010a). *ICCS 2009 International Report. Civic knowledge, attitudes and engagement among lower secondary school students in thirty-eight countries.* Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Schulz, W., Ainley, J., Fraillon, J., Kerr, D. & Losito, B. (2010b). *Initial Findings from the International Civic and Citizenship Education Study.* Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Schulz, W., Ainley, J., Friedman, T. & Lietz, P. (2011). *ICCS 2009 Latin American Report: Civic knowledge and attitudes among lower secondary students in six*

Latin American countries. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Schulz, W., Fraillon, J., Ainley, J., Losito, B., & Kerr, D. (2008) *International civic and citizenship education study. Assessment framework.* Amsterdam: International Association for the Evaluation of Educational Achievement (IEA).

Torney-Purta, J., Lehmann, R., Oswald, H., & Schulz, W. (2001). *Citizenship and education in twenty-eight countries*. Amsterdam: International Association for the Evaluation of Educational Achievement (IEA).

Wu, M. L., Adams, R. J., Wilson, M. R. & Haldane, S. (2007). *ACER ConQuest 2.0: General item response modelling software* [computer program manual]. Camberwell, VIC: ACER Press.

Figure 1: Scatter plot for link item parameter estimates

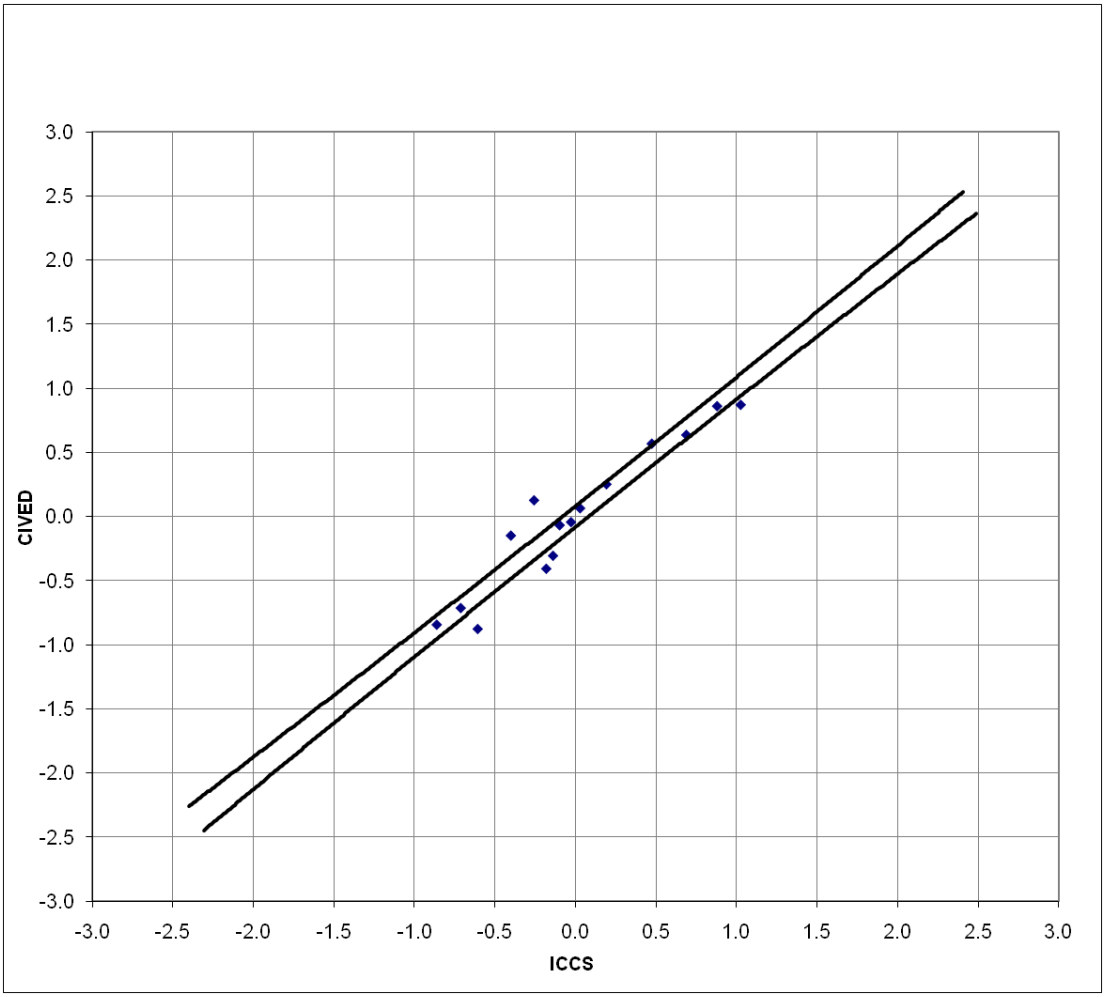


Table 1 **Item parameters and average percent correct for link items**

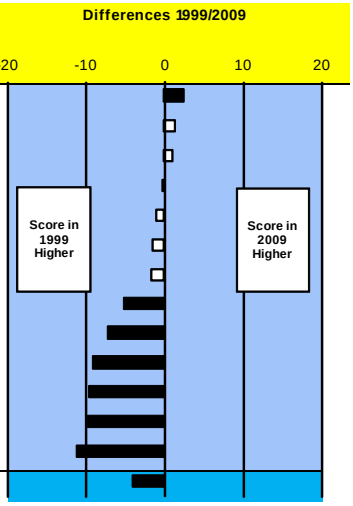
Item	Item parameters		Average percentages of correct responses in: ICCS 2009	CIVED 1999
CI101M1	-0.79	(.02)	72	78
CI104M1	-0.22	(.01)	63	70
CI106M1	-0.62	(.02)	72	79
CI108M1	0.15	(.01)	60	63
CI109M1	-0.32	(.01)	65	71
CI110M1	0.34	(.01)	57	58
CI113M1	-0.06	(.01)	68	67
CI115M1	0.73	(.01)	46	50
CI119M1	0.05	(.01)	61	66
CI120M1	-0.75	(.02)	75	77
CI121M1	0.95	(.01)	43	49
CI127M1	0.96	(.01)	38	47
CI128M1	0.22	(.01)	66	63
CI129M1	0.66	(.01)	51	54
CI130M1	0.02	(.01)	61	64

Table 2 **Test reliabilities for link items (Cronbach's alpha)**

Country	Reliability	Number of items
Bulgaria	0.78	15
Czech Republic	0.77	17
England	0.77	17
Estonia	0.70	15
Finland	0.80	16
Greece	0.82	17
Italy	0.80	17
Latvia	0.69	17
Lithuania	0.73	16
Norway	0.80	17
Poland	0.82	17
Slovak Republic	0.74	14
Slovenia	0.76	16
Sweden	0.78	15
Switzerland	0.70	15
European ICCS median	0.77	16

Table 3 **Changes in civic content knowledge between 1999 and 2009 in European ICCS countries**

Country	Years of schooling	Mean Scale Score 2009	Average age 2009	Mean Scale Score 1999	Average age 1999	Differences between 1999 and 2009
Slovenia	9	104 (0.6)	14.7	102 (0.5)	14.8	3 (1.0)
Finland	8	109 (0.7)	14.7	108 (0.7)	14.8	1 (1.1)
Estonia	8	95 (0.9)	15.0	94 (0.5)	14.7	1 (1.2)
Lithuania	8	94 (0.6)	14.7	94 (0.7)	14.8	0 (1.1)
Italy	8	100 (0.7)	13.8	101 (0.7)	13.9	-1 (1.2)
Latvia	8	91 (0.6)	14.8	92 (0.9)	14.5	-1 (1.2)
Switzerland (German) †	8	94 (1.0)	14.8	95 (0.9)	15.0	-2 (1.5)
Norway †~	9	97 (0.8)	14.7	103 (0.5)	14.8	-5 (1.1)
Greece	9	102 (0.8)	14.7	109 (0.7)	14.7	-7 (1.3)
Poland	8	103 (1.0)	14.9	112 (1.3)	15.0	-9 (1.8)
Slovak Republic[1]	8	97 (1.1)	14.4	107 (0.6)	14.3	-10 (1.4)
Czech Republic †	8	93 (0.5)	14.4	103 (0.8)	14.4	-10 (1.1)
Bulgaria	8	88 (0.9)	14.7	99 (1.1)	14.9	-11 (1.5)
Average		98 (0.0)	14.6	101 (0.0)	14.7	-4 (0.1)



() Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.

† Met ICCS guidelines for sampling paticipation rates only after replacement schools were included.

‡ Nearly satisfied ICCS guidelines for sample participation only after replacement schools were included.

~ In 1999, overall participation rate after replacement less than 75 percent.

[3] National Desired Population does not cover all of International Desired Population.

[2] In 1999, country surveyed the same cohort of students but at the beginning of the next school year.

[3] In 1999, country surveyed the same cohort of students but at the at the beginning of the school year.

Difference statistically significant (p < .05)

Difference not statistically significant.

Table 4 **Gender differences in civic content knowledge in comparison**

	Differences (females-males)					
Country	**CIVED 1999**			**ICCS 2009**		
Bulgaria	1.8	(1.1)		2.5	(1.3)	
Czech Republic	-2.4	(0.8)	*	-0.2	(0.9)	
Estonia	0.7	(0.7)		0.5	(1.3)	
Finland	0.7	(0.9)		1.3	(1.3)	
Greece	0.3	(0.7)		4.2	(1.6)	*
Italy	1.0	(0.6)		0.6	(1.1)	
Latvia	2.9	(0.9)	*	1.5	(1.1)	
Lithuania	1.6	(0.6)	*	2.4	(1.0)	*
Norway	-1.7	(0.8)	*	0.2	(1.2)	
Poland	1.1	(1.4)		1.1	(1.3)	
Slovak Republic	-0.8	(0.7)		1.3	(1.6)	
Slovenia	3.1	(0.8)	*	4.1	(1.2)	*
Switzerland	-3.0	(0.7)	*	-1.4	(1.2)	
European ICCS average	0.4	(0.2)		1.4	(0.3)	*

Statistically significant differences ($p < 0.05$) marked with asterisk.

Table 5 **Differences in civic content knowledge by country of birth in comparison**

	Differences (born in country - others)					
Country	**CIVED 1999**			**ICCS 2009**		
Bulgaria	-1.1	(7.4)		5.5	(5.3)	
Czech Republic	-0.2	(2.9)		1.0	(3.3)	
Estonia	2.9	(1.7)		0.9	(4.0)	
Finland	5.4	(2.6)	*	15.3	(3.9)	*
Greece	4.8	(1.4)	*	9.6	(2.4)	*
Italy	5.2	(2.3)	*	10.5	(2.4)	*
Latvia	4.3	(1.9)	*	3.0	(4.1)	
Lithuania	-1.6	(2.0)		5.8	(4.6)	
Norway	9.9	(1.5)	*	6.3	(2.5)	*
Poland	13.3	(3.8)	*	-0.6	(7.6)	
Slovak Republic	1.4	(3.0)		5.1	(7.5)	
Slovenia	6.0	(1.8)	*	-0.9	(4.1)	
Switzerland	6.7	(1.4)	*	-0.5	(3.3)	
European ICCS average	4.4	(0.8)	*	4.7	(1.3)	*

Statistically significant differences ($p < 0.05$) marked with asterisk.

Table 6 **Differences in civic content knowledge by parental education in comparison**

	Differences (one parent at university - others)					
Country	**CIVED 1999**			**ICCS 2009**		
Bulgaria	8.2	(1.7)	*	9.5	(1.5)	*
Czech Republic	11.1	(1.3)	*	6.0	(1.1)	*
Estonia	8.6	(0.7)	*	4.1	(1.4)	*
Finland	8.1	(1.1)	*	5.8	(1.5)	*
Greece	13.0	(1.0)	*	11.4	(1.5)	*
Italy	10.6	(1.2)	*	8.4	(1.6)	*
Latvia	7.3	(1.0)	*	6.8	(1.1)	*
Lithuania	8.2	(0.9)	*	10.1	(1.4)	*
Norway	10.2	(0.9)	*	8.4	(1.3)	*
Poland	11.1	(1.2)	*	11.0	(1.4)	*
Slovak Republic	9.3	(1.1)	*	8.9	(1.5)	*
Slovenia	12.5	(0.8)	*	8.5	(1.4)	*
Switzerland	6.0	(1.1)	*	9.2	(1.4)	*
European ICCS average	9.5	(0.3)	*	8.3	(0.4)	*

Statistically significant differences ($p < 0.05$) marked with asterisk.

Table 7 **Percentages of missing responses for ethnic group items in comparison**

	CIVED 1999				ICCS 1999			
Country	**Item 1**	**Item 2**	**Item 3**	**Item 4**	**Item 1**	**Item 2**	**Item 3**	**Item 4**
Bulgaria	12	14	13	30	4	4	5	5
Czech Republic	3	4	4	13	1	1	1	1
Estonia	4	7	6	15	2	2	2	2
Finland	5	7	11	20	2	2	3	3
Greece	8	8	6	16	2	2	2	2
Italy	12	12	6	20	1	1	1	1
Latvia	10	10	12	22	1	2	2	2
Lithuania	9	9	8	25	1	1	1	1
Norway	5	11	17	25	5	5	5	6
Poland	5	7	4	15	1	1	1	1
Slovak Republic	2	4	4	10	0	1	1	1
Slovenia	7	8	6	21	2	2	2	2
Switzerland	8	9	11	16	1	1	1	1
European ICCS average	7	8	8	19	2	2	2	2

Table 8 **Percentages of agreement for item 1 and 2 in comparison**

	Item 1						Item 2					
Country	**CIVED**		**ICCS**		**Difference**		**CIVED**		**ICCS**		**Difference**	
Bulgaria	80	(1.5)	90	(0.9)	10	(1.7)	81	(1.3)	86	(0.8)	5	(1.5)
Czech Republic	93	(0.7)	90	(0.5)	-3	(0.8)	91	(0.7)	88	(0.6)	-2	(0.9)
Estonia	94	(0.4)	97	(0.4)	3	(0.6)	87	(0.7)	94	(0.6)	7	(0.9)
Finland	94	(0.6)	93	(0.5)	-2	(0.8)	93	(0.6)	91	(0.5)	-2	(0.8)
Greece	82	(0.8)	92	(0.6)	10	(1.0)	89	(0.7)	85	(0.7)	-3	(1.0)
Italy	88	(0.8)	93	(0.5)	6	(0.9)	85	(0.7)	89	(0.6)	4	(1.0)
Latvia	89	(0.9)	92	(0.8)	3	(1.2)	87	(1.1)	92	(0.8)	5	(1.4)
Lithuania	92	(0.6)	96	(0.4)	3	(0.7)	91	(0.5)	94	(0.5)	3	(0.7)
Norway	93	(0.5)	91	(0.8)	-3	(0.9)	89	(0.8)	88	(0.9)	0	(1.2)
Poland	91	(0.8)	95	(0.4)	4	(0.9)	88	(0.8)	94	(0.5)	6	(1.0)
Slovak Republic	91	(0.9)	94	(0.7)	3	(1.1)	91	(1.0)	94	(0.6)	3	(1.2)
Slovenia	89	(0.8)	92	(0.5)	3	(0.9)	85	(0.8)	90	(0.6)	5	(0.9)
Switzerland	77	(1.4)	92	(0.7)	15	(1.6)	74	(1.5)	90	(1.1)	16	(1.8)
European ICCS average	89	(0.2)	93	(0.2)	4	(0.3)	87	(0.3)	91	(0.2)	4	(0.3)

Table 9 **Percentages of agreement for item 3 and 4 in comparison**

	Item 3						Item 4					
Country	**CIVED**		**ICCS**		**Difference**		**CIVED**		**ICCS**		**Difference**	
Bulgaria	79	(1.6)	87	(0.8)	7	(1.8)	62	(1.6)	57	(1.1)	-6	(1.9)
Czech Republic	87	(0.9)	86	(0.5)	-1	(1.0)	72	(1.1)	57	(0.9)	-15	(1.5)
Estonia	88	(0.6)	90	(0.8)	2	(1.0)	75	(0.9)	78	(0.9)	4	(1.3)
Finland	81	(1.1)	81	(0.8)	0	(1.3)	83	(0.8)	70	(1.1)	-13	(1.4)
Greece	91	(0.7)	91	(0.8)	1	(1.1)	75	(0.8)	62	(1.1)	-13	(1.4)
Italy	89	(0.7)	92	(0.7)	2	(1.0)	72	(1.0)	66	(1.1)	-5	(1.5)
Latvia	82	(1.3)	83	(1.2)	0	(1.7)	57	(1.5)	58	(1.5)	1	(2.1)
Lithuania	90	(0.7)	93	(0.6)	3	(0.9)	74	(0.9)	67	(1.0)	-7	(1.3)
Norway	79	(1.0)	91	(0.8)	12	(1.3)	81	(0.9)	80	(1.2)	-1	(1.5)
Poland	90	(0.6)	92	(0.6)	2	(0.8)	68	(1.1)	78	(0.9)	10	(1.4)
Slovak Republic	82	(1.2)	87	(0.8)	5	(1.4)	72	(1.3)	71	(1.0)	-1	(1.6)
Slovenia	87	(0.7)	88	(0.7)	1	(1.0)	58	(1.0)	65	(1.0)	7	(1.5)
Switzerland	66	(1.7)	87	(1.1)	21	(2.1)	71	(1.6)	66	(1.3)	-5	(2.1)
European ICCS average	84	(0.3)	88	(0.2)	4	(0.4)	71	(0.3)	67	(0.3)	-3	(0.4)