

**Knowledge of
Content and Skills
in Interpreting
Civic Information**

HIGHLIGHTS RELATING TO CIVIC KNOWLEDGE

- The results show that it is possible to construct a meaningful, reliable and valid international test of student knowledge about democratic institutions, principles, processes and related topics despite differences in the political systems in different participating countries.
- The differences between countries in mean performance on this test are in general not large. Twenty-five of the 28 countries differ by less than half a standard deviation from the international average.
- Students' abilities to answer questions requiring knowledge of civic content and questions requiring skill in *interpreting* civic-related material are distinguishable empirically. Although this distinction between content knowledge and skills in interpretation has a very limited influence on the countries' rankings, it adds to the understanding of countries' specific strengths and weaknesses.
- Unlike studies of earlier decades, this study reveals no significant differences in mean performance between boys and girls in 27 of the 28 countries, when the comparison is made without controlling for other variables.
- Students reporting more home literacy resources consistently do better on the test.
- Over 75 percent of the students in most participating countries are able to answer questions dealing with the fundamental nature of laws and political rights. However, few students in the participating countries can answer more demanding questions on the test that have to do, for example, with deciding between election candidates based on their policy positions, understanding processes of political reform, and grasping the implications of economic and political choices made by policy-makers.

The IEA Civic Education Study faced considerable skepticism about the possibility of developing a valid test to measure civic knowledge across diverse political systems. Chapter 2 has described the process by which we met this challenge, namely producing a test rooted in the content domains defined collaboratively in Phase 1 of the study, meeting IEA technical standards and allowing for extensive procedures of quality control. For a review of the previous research in this area, see Panel 3.1

In this chapter, we present a summary of the results achieved on this test by nearly 90,000 students who constituted nationally representative samples from the 28 participating countries. We describe the methods applied in the scoring and scaling of student responses, and then display the distribution of test scores by country, followed by an analysis of how the test can be partitioned into two dimensions relating to knowledge of content and skills in interpreting civic-related information. Finally, we deal with the ways in which the key variables of gender and home literacy resources are related to civic knowledge. In short, the chapter provides a base for the comparative assessment of what students in participating countries know about the nature and workings of democracy and, to some extent, about the other two core content domains of the study.

PANEL 3.1 Previous Research on Students' Civic Knowledge

First, it is important to acknowledge the role of general literacy in acquiring political knowledge. For example, Chall and Henry (1991) noted that considerably more than a minimal level of literacy is required for understanding documents such as constitutions or for locating information in sources such as newspapers.

As part of a more specific look at the role of knowledge in the context of civic education, the 1971 IEA Civic Education Study used a test of 47 items for 14-year-olds (Torney, Oppenheim & Farnen, 1975). In this study, students in the Federal Republic of Germany had the second highest cognitive score. The United States ranked fourth, Finland ranked fifth, and Italy ranked seventh (out of eight countries). Among other countries that participated in the 1971 but not in the 1999 study, students from the Netherlands ranked first, those from Israel ranked third, those from New Zealand ranked sixth, and those from Ireland ranked eighth. At age 14, males performed higher on the test than females in the Federal Republic of Germany, Finland and the United States (but not in Italy) in a comparison similar to that reported in this chapter. Those from higher socioeconomic backgrounds performed at a higher level in all the countries, although the differences were especially large in the United States (Torney *et al.*, 1975, pp.138, 156). The encouragement of independent expression of opinion in the classroom was a positive predictor in all the countries (p.140).

In some countries there have also been large-scale national assessments of civic knowledge. In the United States a National Assessment of Educational Progress (NAEP) regularly tests students at Grades 4, 8 and 12 (ages approximately 9, 13 and 17) in civic-related content areas. Multiple choice items are used, along with items that require students to write a response (sometimes relating to a picture of a historical event, cartoon or newspaper article). No attempt is made to separate performance on knowledge of content and skills in interpretation (as in the current IEA study). Proficiency levels for the 'total knowledge' scores are set by experts. Most students are classified as having 'basic' rather than 'proficient' or 'advanced' understanding (Lutkus, Weiss, Campbell, Mazzeo & Lazer, 1999; Torney-Purta, 2000). There is a substantial gap between the scores of students from more and less affluent and highly and less educated home backgrounds (Niemi & Junn, 1998).

Gender differences in the NAEP have been variable. In the 1988 assessment, males tended to perform at a somewhat higher level than females. Gender differences were especially pronounced in knowledge of political parties, elections and protest activities. In the 1998 assessment, these differences were either very small or showed females to have slight superiority. The most comprehensive recent analysis of school-based predictors of achievement in NAEP, by Niemi and Junn (1998, using the 1988 data), found that frequent testing seemed to be counterproductive in students' learning of civic content. The taking of classes in which civic topics were studied and where participation in role-playing elections or mock trials was included seemed to have a positive effect.

A study in Australia with 1,000 students from Years 5 and 9 tapped political understanding by asking questions to which students wrote answers (Doig, Piper, Mellor & Masters, 1993/94). Topics included the meaning and origin of laws, the electoral process, influences on political decisions, processes of enactment and implementation of parliamentary decisions, and the meaning of democracy. Each response was scored as being at one of several levels ranging from simplistic, vague or confused to sophisticated and complex (including the ability to apply principles). The average student was found to be able to 'recognize key aspects of democracy in a generalized way'. On average, Year 9 students had higher scores than Year 5 students. Females at Year 5 had higher scores than males, but there were no significant gender differences among the older students.

For a fuller review of studies in these areas, see Torney-Purta, Hahn and Amadeo (2001).

CIVIC KNOWLEDGE IN THE 1999 IEA INSTRUMENT: HOW IT WAS MEASURED

The IEA civic knowledge test consists of 38 items, 25 of which refer to knowledge of content (Type 1) and 13 to skills in interpretation (Type 2). All items were given in a multiple-choice format, with student responses coded as correct or incorrect. The items cover a broad range of content areas selected from a much larger set of trial items after intensive piloting (see Chapter 2). In the final test version, the international average of correct answers was 64 percent, which indicates that, for the majority of students, the test was far from being too difficult. The results of Rasch scaling as well as classical item statistics (Appendix Tables C.1 and C.2) show that the test has satisfactory characteristics; for example, alpha coefficients are at least .85 in all countries. More detail will appear in the technical report (Lehmann *et al.*, forthcoming).

SAMPLE ITEMS AND THEIR RESULTS

To convey a notion of what was measured by the test, five examples are given and briefly discussed here (refer to Figures 3.1a-e).

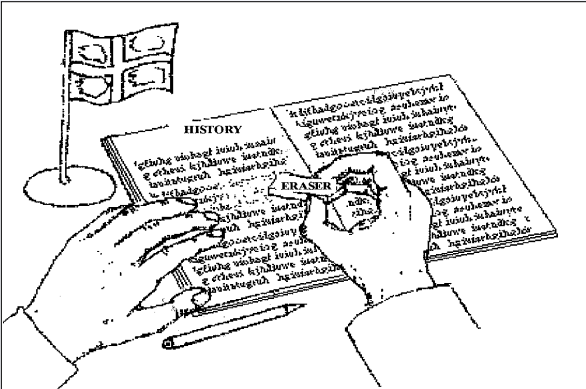
Figure 3.1a Item Example: Identify a non-democratic government

Country	Correct Answers (in %)	Example 3 (Item #17) Type 1: Knowledge of Content
Australia	50 (1.3)	<p>17. Which of the following is most likely to cause a government to be called non-democratic?</p> <p>A. People are prevented from criticising the government.*</p> <p>B. The political parties criticise each other often.</p> <p>C. People must pay very high taxes.</p> <p>D. Every citizen has the right to a job.</p>
Belgium (French)	51 (1.8)	
Bulgaria	53 (2.0)	
Chile	44 (1.3)	
Colombia	38 (1.9)	
Cyprus	59 (1.3)	
Czech Republic	60 (1.6)	
Denmark	46 (1.1)	
England	45 (1.1)	
Estonia	39 (1.4)	
Finland	63 (1.3)	
Germany	56 (1.2)	
Greece	67 (1.2)	
Hong Kong (SAR)	73 (1.3)	
Hungary	45 (1.2)	
Italy	63 (1.5)	
Latvia	36 (1.9)	
Lithuania	44 (1.6)	
Norway	57 (1.0)	
Poland	65 (2.3)	
Portugal	55 (1.5)	
Romania	42 (1.8)	
Russian Federation	57 (2.3)	
Slovak Republic	60 (1.6)	
Slovenia	50 (1.3)	
Sweden	66 (1.6)	
Switzerland	56 (1.6)	
United States	53 (1.7)	
International Sample	53 (0.3)	

() Standard errors appear in parentheses.
* Correct answer.

Source: IEA Civic Education Study, Standard Population of 14-year-olds tested in 1999.

Figure 3.1b Item Example: This is the way history textbooks are sometimes written

Country	Correct Answers (in %)	Example 5 (Item #36) Type 2: Skills in Interpretation
Australia	75 (1.2)	 <p>36. What is the message or main point of this cartoon? History textbooks ...</p> <p>A. are sometimes changed to avoid mentioning problematic events from the past.*</p> <p>B. for children must be shorter than books written for adults.</p> <p>C. are full of information that is not interesting.</p> <p>D. should be written using a computer and not a pencil.</p>
Belgium (French)	66 (2.1)	
Bulgaria	47 (2.3)	
Chile	49 (1.5)	
Colombia	48 (2.3)	
Cyprus	53 (1.1)	
Czech Republic	54 (1.5)	
Denmark	60 (1.0)	
England	76 (1.2)	
Estonia	39 (1.2)	
Finland	65 (1.3)	
Germany	61 (0.9)	
Greece	56 (1.3)	
Hong Kong (SAR)	76 (1.4)	
Hungary	67 (1.3)	
Italy	61 (1.3)	
Latvia	48 (1.7)	
Lithuania	48 (1.4)	
Norway	49 (1.0)	
Poland	64 (2.1)	
Portugal	49 (1.1)	
Romania	26 (1.7)	
Russian Federation	45 (2.1)	
Slovak Republic	72 (1.5)	
Slovenia	56 (1.1)	
Sweden	52 (1.2)	
Switzerland	67 (1.4)	
United States	79 (1.4)	
International Sample	57 (0.3)	

() Standard errors appear in parentheses.
* Correct answer.

Source: IEA Civic Education Study, Standard Population of 14-year-olds tested in 1999.

The first example (Figure 3.1a; see also example 3 in Figure 3.2) is a typical Type 1 item. It requires respondents to demonstrate knowledge of content by identifying a practice that ‘most likely causes a government to be called non-democratic’. Figure 3.1a displays the question and the four answers from which the students had to choose. Among these four answers, the first one (A: ‘People are prevented from criticising the government’) is the correct response. The percentages of correct answers range from 36 to 73 percent. The average of correct answers across all countries (equally weighted) is 53 percent. If this international mean is compared with the overall percentage of correct answers in the test (64 percent), it is clear that the item is relatively difficult, although certainly within reach for most of the students. The correct answer requires a reliable knowledge base as to the properties of democratic governments and the ability to apply that knowledge to the opposite case (‘non-democratic government’).

The second example (Figure 3.1b; example 5 in Figure 3.2) is a Type 2 item that is intended to measure skills in the interpretation of civic-related material. Figure 3.1b demonstrates how the item was presented. There is a cartoon showing someone erasing words from a book, presumably one on the history of a nation as suggested by a flag and the word 'history'. The students were asked about the message or main point of this cartoon and had to select the correct response (A: 'History textbooks are sometimes changed to avoid mentioning problematic events from the past'). The distribution of correct answers across countries ranges from 26 to 79 percent. The international mean of 57 percent correct answers shows that this item is less difficult for students than the previous example, although it is slightly harder than the test on average. The task relates to the domain of national identity and international relations and requires the ability to interpret or comprehend the message that the cartoonist has attempted to convey.

Figure 3.1c Item Example: This is an election leaflet

Country	Correct Answers (in %)	Example 6 (Item #23) Type 2: Skills in Interpretation
Australia	78 (1.3)	<div style="border: 1px solid black; padding: 10px; text-align: center;"> <p>We citizens have had enough! A vote for the Silver Party means a vote for higher taxes. It means an end to economic growth and a waste of our nation's resources. Vote instead for economic growth and free enterprise. Vote for more money left in everyone's wallet! Let's not waste another 4 years! VOTE FOR THE GOLD PARTY.</p> </div> <p>23. This is an election leaflet which has probably been issued by ...</p> <p>A. the Silver Party. B. a party or group in opposition to the Silver Party.* C. a group which tries to be sure elections are fair. D. the Silver Party and the Gold Party together.</p>
Belgium (French)	56 (1.8)	
Bulgaria	47 (2.4)	
Chile	54 (1.5)	
Colombia	40 (2.4)	
Cyprus	81 (0.9)	
Czech Republic	66 (1.6)	
Denmark	49 (1.1)	
England	75 (1.2)	
Estonia	54 (1.4)	
Finland	85 (0.8)	
Germany	81 (0.9)	
Greece	73 (1.3)	
Hong Kong (SAR)	76 (1.4)	
Hungary	78 (1.2)	
Italy	85 (1.2)	
Latvia	44 (1.9)	
Lithuania	55 (1.6)	
Norway	57 (0.9)	
Poland	58 (2.0)	
Portugal	55 (1.3)	
Romania	46 (2.0)	
Russian Federation	45 (1.9)	
Slovak Republic	66 (1.6)	
Slovenia	75 (1.0)	
Sweden	73 (1.5)	
Switzerland	77 (1.3)	
United States	83 (1.4)	
International Sample	65 (0.3)	

() Standard errors appear in parentheses.
* Correct answer.

Source: IEA Civic Education Study, Standard Population of 14-year-olds tested in 1999.

Figure 3.1d Item Example: Importance of many organisations for democracy

Country	Correct Answers (in %)	Example 7 (Item #07) Type 1: Knowledge of Content
Australia	78 (1.2)	<p>7. In a democratic country [society] having many organisations for people to join is important because this provides ...</p> <p>A. a group to defend members who are arrested.</p> <p>B. many sources of taxes for the government.</p> <p>C. opportunities to express different points of view.*</p> <p>D. a way for the government to tell people about new laws.</p>
Belgium (French)	68 (1.6)	
Bulgaria	71 (1.9)	
Chile	69 (1.1)	
Colombia	60 (2.0)	
Cyprus	80 (1.1)	
Czech Republic	76 (1.2)	
Denmark	75 (0.9)	
England	79 (1.0)	
Estonia	61 (1.1)	
Finland	82 (1.0)	
Germany	67 (1.0)	
Greece	76 (0.9)	
Hong Kong (SAR)	79 (1.1)	
Hungary	46 (1.3)	
Italy	71 (1.4)	
Latvia	55 (1.8)	
Lithuania	61 (1.4)	
Norway	69 (0.9)	
Poland	78 (1.6)	
Portugal	59 (1.2)	
Romania	48 (2.2)	
Russian Federation	68 (1.6)	
Slovak Republic	75 (1.1)	
Slovenia	62 (1.1)	
Sweden	70 (1.5)	
Switzerland	68 (1.3)	
United States	78 (1.4)	
International Sample	69 (0.3)	
<p>() Standard errors appear in parentheses.</p> <p>* Correct answer.</p>		

Source: IEA Civic Education Study, Standard Population of 14-year-olds tested in 1999.

The third example (Figure 3.1c; example 6 in Figure 3.2) is also a Type 2 item, in this case referring to institutions and practices in democracy. Here, students were asked to interpret an electoral leaflet directed against an imaginary party (presumably in power) and to indicate which political group had probably issued it. The correct answer is B ('a party or group in opposition to the Silver Party'). The lowest proportion of correct answers found in any country is 40 percent and the highest 85 percent. The international average is 65 percent. Thus, the item is slightly easier than the international average of the test as a whole. There are several clues suggesting the correct solution, although the interpretation of some of these requires quite complex inferences as to the two mentioned parties' approaches to taxation and government spending. The last line of the leaflet ('Vote for the Gold Party') is unambiguous and clearly marks its origin. As such, the interpretative task is primarily to identify the alleged negative economic consequences of the Silver Party's fiscal policies and to recognize that such arguments could come only from an opposing group.

The fourth example (Figure 3.1d; example 7 in Figure 3.2) is a Type 1 item, aiming at civic-related content knowledge, more specifically knowledge as to the rights and duties of citizens in a democratic country. The students could choose from four reasons justifying ‘freedom of association’, with the correct response being C (‘Having many organizations for people to join is important because this provides “opportunities to express different points of view”’). National percentages of correct responses range from 46 to 82 percent. With an international mean of 69 percent correct, this item is clearly among the easier ones in the test. It calls for basic background knowledge, or perhaps some recall of politics-related experience, to rule out the incorrect responses and to select the correct one.

Figure 3.1e Item Example: Function of having more than one political party

Country	Correct Answers (in %)	Example 8 (Item #11) Type 1: Knowledge of Content
Australia	75 (1.3)	<p>11. In democratic countries what is the function of having more than one political party?</p> <p>A. To represent different opinions [interests] in the national legislature [e.g. Parliament, Congress].*</p> <p>B. To limit political corruption.</p> <p>C. To prevent political demonstrations.</p> <p>D. To encourage economic competition.</p>
Belgium (French)	67 (1.7)	
Bulgaria	70 (1.6)	
Chile	60 (1.2)	
Colombia	54 (1.6)	
Cyprus	88 (0.9)	
Czech Republic	79 (1.0)	
Denmark	84 (0.8)	
England	78 (1.0)	
Estonia	62 (1.2)	
Finland	80 (1.0)	
Germany	84 (0.9)	
Greece	85 (0.7)	
Hong Kong (SAR)	76 (1.1)	
Hungary	75 (1.2)	
Italy	86 (0.9)	
Latvia	57 (1.7)	
Lithuania	68 (1.2)	
Norway	83 (0.7)	
Poland	82 (1.1)	
Portugal	84 (0.8)	
Romania	67 (1.7)	
Russian Federation	71 (1.6)	
Slovak Republic	77 (1.0)	
Slovenia	81 (0.7)	
Sweden	75 (1.5)	
Switzerland	82 (0.9)	
United States	72 (1.5)	
International Sample	75 (0.2)	

() Standard errors appear in parentheses.
* Correct answer.

Source: IEA Civic Education Study, Standard Population of 14-year-olds tested in 1999.

The fifth example (Figure 3.1e; example 8 in Figure 3.2) is another Type 1 item, this time relating to institutions and practices in a democracy. In this item, four potential functions of a political system with more than one party were presented to the students who had to select the correct one (A: ‘to represent different opinions [interests] in the national legislature’). The task turned out to be quite easy, with an international average percentage correct of 75 and a range across countries of 54 to 88 percent. In terms of its cognitive demands, this item is quite similar to the previous example, that is, a certain amount of political background knowledge and/or politics-related experience is needed if the correct response is to be identified. Some might argue that option B (‘to limit political corruption’) is not entirely wrong. The students who generally did well on the test, however, shared the conviction that the function of parties to represent different opinions or interests refers to a more fundamental role and is the more appropriate choice.

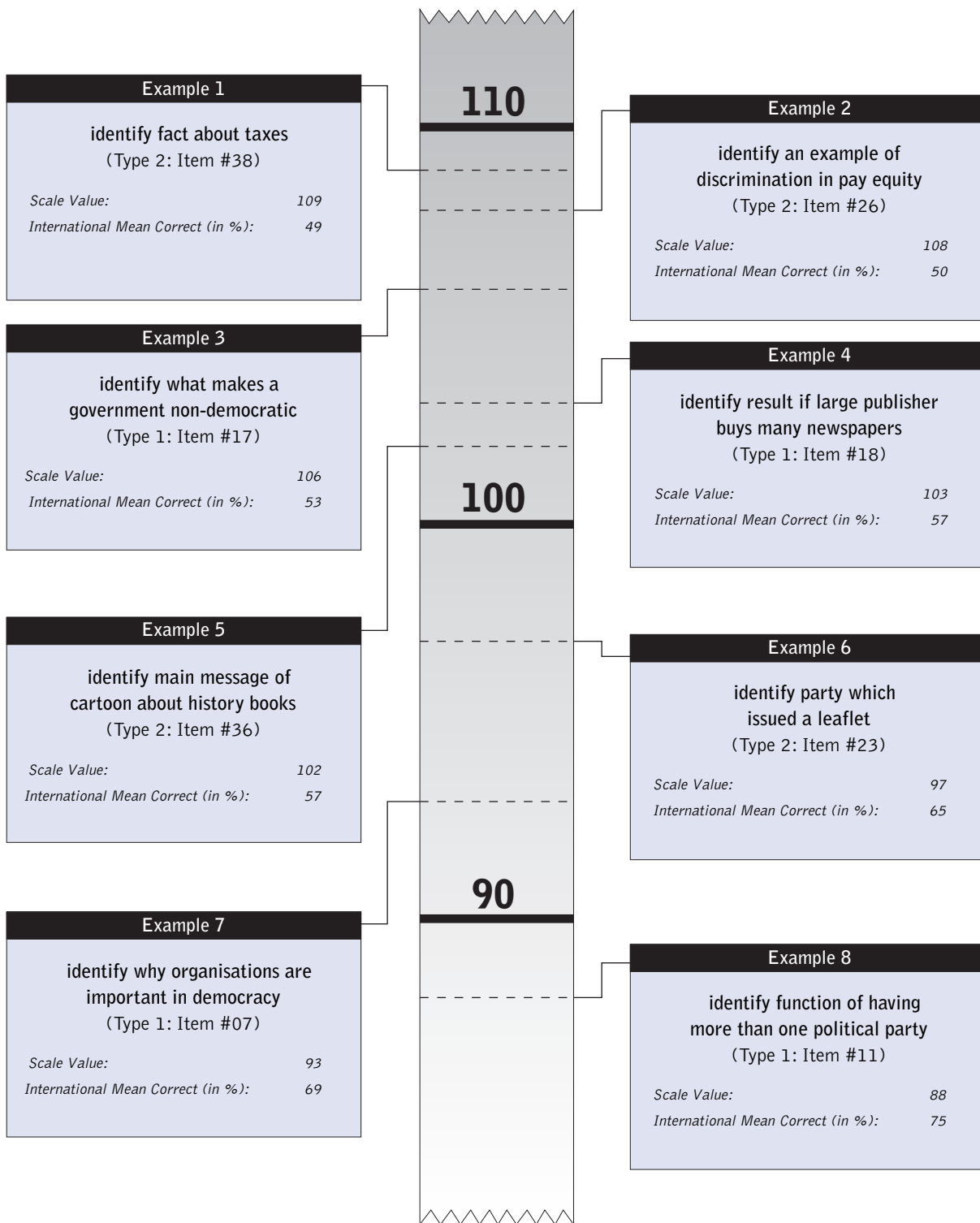
Three more sample items from the IEA test are given in Appendix A (Figures A.1a–c). They are intended to illustrate further the kind of questions and cognitive demands presented in the questions. Table A.1 in Appendix A lists the domain categories and short titles for all items contained in the test, along with the percentages of correct responses (international means).

ESTIMATION OF ITEM DIFFICULTIES AND STUDENT ABILITIES

One can always rank the items of a test according to the respective percentages of correct responses, the easiest ones being those with the highest percentage of correct answers and the most difficult ones those with the lowest. These percentages, however, are not the most informative measures of difficulty, since they do not take into account the ability levels of those who answered the item correctly. Similarly, ‘percent correct’ (or the sum of correct responses) is not a very good measure of student ability, because it assumes implicitly that all items are equally difficult.

When certain assumptions are met, it is possible to apply modern scaling techniques to arrive at an interval scale onto which measures of both item difficulty and student ability can be projected at the same time. The most important of these conditions is that a student with an ability thus determined solves—at a *defined or reasonable level of success*—all of the items with difficulties up to his or her ability level and fails most of the harder items. Following the example of the Third International Mathematics and Science Study (TIMSS), we chose in the present study a probability of .65 to represent this defined level of success. Because the simultaneous estimation of student abilities and item difficulties is based on probability functions, such an approach is sometimes called ‘probabilistic’. The more widely used term is ‘Item Response Theory’ or ‘IRT’ (see, for example, Hambleton, Swaminathan & Rogers, 1991). We used one of its models (the so-called one-parameter model that produces maximum likelihood estimates) to scale the 38 cognitive test items, once we had established that the required underlying assumptions held empirically. Because this model leaves researchers free to choose the mean and the standard deviation for the metric to be used, we set the international mean of the scale for civic knowledge (and two sub-scales, to be discussed later) to 100, with a standard deviation of 20.

Figure 3.2 International Difficulty Map for Sample Items Relating to Civic Knowledge



NOTE: Each item was placed onto the International Civic Knowledge Scale. Items are shown at the point on the scale where students with that level of proficiency had a 65 percent probability of providing a correct response.

Because percentages and scale values are rounded to the nearest whole number, some results may appear inconsistent.

Source: IEA Civic Education Study, Standard Population of 14-year-olds tested in 1999.

The item difficulty estimates produced by using an IRT model allowed us to demonstrate the substantive meaning of the items in the context of the test as a whole. Figure 3.2 provides an ‘item difficulty map’ that illustrates the relationship between the item-specific performance levels and the international knowledge test score as defined.

In this figure, the item examples (including those given in Appendix A) are placed on the scale—the gray bar in the middle—at the point where a student with the respective ability has a probability of .65 to provide the correct answer. For example, a student with a scale score of at least 106 on the civic knowledge scale has a chance of two in three or better to identify correctly what makes a government non-democratic (item example 3). Similarly, students with scale scores lower than 97 will have less than a 65 percent chance to identify correctly the origin of the imaginary electoral leaflet (item example 6).¹

We had determined that the probability level of .65, which links item difficulties and student abilities, should be set as a formal characteristic of the metric used. Substantively, however, we considered this to be an appropriate interpretation of the somewhat vague term ‘reasonable level of success’. It reflects our attempt to take into account knowledge expectations with a higher than 50 percent chance, even though this level is still below that which some might call mastery of the subject.

Figure 3.2 also displays, for each item example, the international average of correct answers. For the test as a whole, including those items not illustrated here, the percentages of correct responses range from 35 percent for the hardest item in the test to 85 percent for the easiest one. The respective difficulty parameters, or scale values, are 77 for the easiest item and 121 for the hardest. A fairly wide range of difficulties and student abilities therefore could be covered by the test. Table A.1 in Appendix A also contains the difficulty parameters for each item. Thus, it is possible to see how item difficulties are distributed across content domains.

One crucial point in the selection of items was to ensure that the item difficulties, and consequently the estimated student abilities, were truly comparable across countries. We had anticipated that this assumption would be difficult to meet, that is, ‘differential item-functioning’ or ‘item-by-country interaction’ would occur. This is the case when an item of a certain international difficulty level is relatively easier or harder for students in a country than one would expect on the basis of that country’s overall mean. Different civic education curricula or differences in the historical and political context might cause such deviations in a country’s response pattern from the international findings. Although there were 1,064 item-by-country pairs (38 items for each of 28 countries), only eight of these showed a clear item-by-country interaction. Because there were so few, we decided not to re-estimate (‘float’) the item difficulty parameters for the countries concerned, but to rely on the fact that the interaction effects within a country sum to zero.

Another potential problem was that, in a particular country, some items might not discriminate well between high- and low-achieving students. This happens when many able students in that country fail to choose the correct response, or when many weaker students do choose it (‘item misfit’). Again, we found that such deviations from the international test characteristics were rare: we encountered 33

instances out of the 1,064. To ensure full international comparability, however, we deemed it appropriate to exclude the misfitting items countrywise from the final analysis of scores by treating them statistically as not having been administered. We therefore re-estimated all parameters, taking these cases into account. In no case is a country's civic knowledge score based on fewer than 35 out of the 38 available items, and in no case did we implicitly penalize a country for not being scored on the full set of 38 items. For further detail, see the technical report (Lehmann *et al.*, forthcoming).

In general, careful analysis showed us that it was possible to construct a meaningful, reliable and valid international test of student knowledge about democratic institutions, principles, processes and related topics that has a high degree of comparability across countries.

CIVIC KNOWLEDGE ACROSS COUNTRIES

Figure 3.3 presents a summary of the results of the international test on civic knowledge for all participating countries. To aid interpretation of these results, we added the date of testing, the tested grade and the average age of students to the figure. For some countries, readers should also take into account the specific information contained in the footnotes and the fuller elaboration given in Chapter 2.

On the basis of the scaling technique just explained, the table within Figure 3.3 contains the average test score and the standard error of sampling for each participating country. The international mean of the distribution (based on the 28 equally weighted national samples) and its standard error are included to give an orientation mark for each country to compare itself to the international average. Thus, the table indicates which countries differ significantly (after correction for multiple comparisons) from the international average.

Ten countries have means that are significantly higher and eight countries have means that are significantly lower than the international mean. The remaining ten countries belong to a middle group with country means that do not differ significantly from the international mean. For the most part, differences between countries within the three major groups are not significant (Figure 3.4).

Poland, the Czech Republic and the Slovak Republic, three post-communist countries from Central and Eastern Europe, belong to the top group. Other countries in this group are the Greek-speaking countries Greece and Cyprus, the United States, Italy and two Nordic countries (Finland and Norway). Students from Hong Kong (SAR) also perform significantly better than the international average.

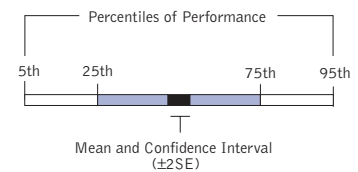
In the group of countries with means significantly below average, only Portugal and French-speaking Belgium are in Western Europe, and both tested very young students. Romania and the three Baltic countries Estonia, Latvia and Lithuania have average scores between 92 and 94. Chile and Colombia, the two Latin American countries in this study, have the lowest means (88 and 86 respectively). Here, 75 percent of the students have scores below the international average.

Figure 3.3 Distributions of Civic Knowledge

Country	Mean Scale Score	Testing Date	Tested Grade*	Mean Age**	Cognitive Civic Competence Scale Score	
					40 60 80 100 120 140 160	
Poland	▲ 111 (1.7)	5/99 - 6/99	8	15.0		
Finland	▲ 109 (0.7)	4/99	8	14.8		
Cyprus	▲ 108 (0.5)	5/99	9	14.8		
Greece	▲ 108 (0.8)	3/99 - 6/99	9	14.7		
Hong Kong (SAR) ³	▲ 107 (1.1)	6/99 - 7/99	9	15.3		
United States ¹	▲ 106 (1.2)	10/99	9	14.7		
Italy	▲ 105 (0.8)	4/99 - 5/99	9	15.0		
Slovak Republic	▲ 105 (0.7)	5/99 - 6/99	8	14.3		
Norway ⁴	▲ 103 (0.5)	4/99 - 6/99	8	14.8		
Czech Republic	▲ 103 (0.8)	4/99 - 5/99	8	14.4		
Australia	● 102 (0.8)	8/99	9	14.6		
Hungary	● 102 (0.6)	3/99	8	14.4		
Slovenia	● 101 (0.5)	4/99	8	14.8		
Denmark ⁴	● 100 (0.5)	4/99	8	14.8		
International sample	● 100 (0.2)	3/99 - 12/99	8/9	14.7		
Germany ²	● 100 (0.5)	4/99 - 7/99	8	n.a.		
Russian Federation ³	● 100 (1.3)	4/99 - 5/99	9	15.1		
England ¹	● 99 (0.6)	11/99	9	14.7		
Sweden ¹	● 99 (0.8)	10/99 - 12/99	8	14.3		
Switzerland	● 98 (0.8)	4/99 - 7/99	8/9	15.0		
Bulgaria	● 98 (1.3)	5/99 - 6/99	8	14.9		
Portugal ⁵	▼ 96 (0.7)	4/99	8	14.5		
Belgium (French) ⁴	▼ 95 (0.9)	3/99 - 4/99	8	14.1		
Estonia	▼ 94 (0.5)	4/99	8	14.7		
Lithuania	▼ 94 (0.7)	5/99	8	14.8		
Romania	▼ 92 (0.9)	5/99	8	14.8		
Latvia	▼ 92 (0.9)	4/99 - 5/99	8	14.5		
Chile	▼ 88 (0.7)	10/99	8	14.3		
Colombia	▼ 86 (0.9)	4/99 and 10/99	8	14.6		

() Standard errors appear in parentheses.

- ▲ Country mean significantly higher than international mean.
- No statistically significant difference between country mean and international mean.
- ▼ Country mean significantly lower than international mean.



- 1 Countries with testing date at beginning of school year.
- 2 National Desired Population does not cover all International Desired Population.
- 3 Countries did not meet age/grade specification.
- 4 Countries' overall participation rate after replacement less than 85 percent.
- 5 In Portugal, Grade 8 selected instead of Grade 9 due to average age. Mean scale score for Grade 9 was 106.

* In Switzerland, Grade 8 was tested mainly in German cantons; Grade 9 mainly in French and Italian cantons. In the Russian Federation, students in Grade 9 have eight or nine years of schooling depending on the duration of the primary school they finished. In 1999 about 70 percent of Russian students tested had eight years of schooling at the end of Grade 9.

** Information on age was not available for Germany. International mean age based on 27 countries only.

Source: IEA Civic Education Study, Standard Population of 14-year-olds tested in 1999.

A closer look at these averages reveals the necessity to consider mean age differences. At the country level, the national average of civic knowledge is correlated with the mean age of the sample ($r = .37$). Thus, countries with an older sample have an advantage over those that tested younger students. While it would be possible to adjust national averages of knowledge for differences in mean age (on the assumption of equal growth with age across all countries), such adjustment changes relatively little in the rank order of countries: the correlation between adjusted and unadjusted country means is $r = .92$.² The issue of age and growth will be explored more fully in a future report.

It is beyond the scope of the present volume to try and explain these country differences. Apart from the mean age of the sample, however, other factors can be shown to be correlated, such as the economic situation (gross national product per capita, $r = .32$, or the general level of literacy, $r = .26$; see Tables 1.1 and 1.2). But again, these findings should be taken primarily as evidence that more thorough analyses are needed, both at the international level and the country level with respect to individual and classroom-level processes.

The graphic representation on the right-hand side of Figure 3.3 shows the distribution of civic knowledge in the tested grade for each country. The cognitive score is shown for the mean as well as for the 5th, 25th, 75th and 95th percentiles. Each percentile point indicates the percentage performing below and up to the respective score. To give an example, 25 percent of the students perform below and up to the score marked by the 25th percentile. Seventy-five percent have attained a result above the corresponding score. The dark boxes in the center of each distribution stand for the country means and their 95 percent confidence interval (mean \pm two standard errors of sampling). For an approach that gives substantive meaning to these percentiles, see Panel 3.2.

The variation of means between countries gives further information about the characteristics of this subject area. Twenty-five of the 28 countries differ by less than half of a standard deviation from the international average. Exceptions are Poland, with the highest national average score of 111, and Chile and Colombia, with mean scores below 90. This range can be compared to the results of other international studies on educational achievement. First, it is similar to that of the IEA Reading Literacy Study (Elley, 1994). Fifteen of the 28 countries participating in the Civic Education Study also participated in that study. The range between the highest and the lowest ranking country was .79 of an international standard deviation, compared with .70 in the present study. Secondly, 22 of the 28 countries also participated in the Third International Mathematics and Science Study (Beaton, Mullis, Martin, Gonzalez, Kelly & Smith, 1996). Among these countries, country means for mathematics achievement differed by 1.79 standard deviations between the highest and the lowest achieving country (and still by 1.10 standard deviations if the special case of Colombia was excluded). One of the reasons for this difference between the Reading Literacy Study and the Civic Education Study on the one hand and TIMSS on the other may be that reading (at the age of 14) and civic knowledge are less closely linked to curriculum and instruction than is mathematics.

PANEL 3.2 How to Interpret the Results for the International Distribution of Civic Knowledge

It is informative to attribute substantive meaning to characteristics of the distribution, such as means and percentiles, by relating them to the item-difficulty parameters. These are indicated in the item examples given and in the item overview presented in Table A.1, Appendix A. This procedure can be done for each country separately. The following provides a demonstration of this mode of interpretation for the results for the weighted international file.

The 5th percentile in the international sample corresponds to a scale value of 71, which is lower than the item difficulty parameter of the easiest item in the test. This is Item 16, scale value 77, which pertains to the domain of national identity and international relations and which asks students to identify the major purpose of the United Nations (Table A.1, Appendix A). Thus, the probability that the lowest achieving students—internationally speaking—will respond correctly to this task is less than .65. In contrast, the probability that some students among the lowest 5 percent in a number of countries (for example, Cyprus, the Slovak Republic and Finland) will be able to identify the main purpose of the United Nations at the defined ‘reasonable level of success’ is .65 or higher.

Similarly, the 25th percentile in the international score distribution corresponds to a scale value of 85. This value is close to the item difficulty parameter (scale value 88) of the example given in the text (Figure 3.1e) in which students were asked to identify the function of having more than one political party in a democratic country. More than 75 percent of the students in the participating countries have a probability of .65 or higher of being able to respond correctly to this item, which refers to one of the fundamental traits of representative democracies. Twenty-five percent of the students in the international sample were found to lack such a level of civic knowledge as exemplified by this item (and others of equal difficulty). Another item of approximately this difficulty level is, for instance, Item 12, which asks who ought to govern in a democracy (Appendix Table A.1).

The international mean of 100 corresponds to an item also pertaining to the domain of democracy and its defining characteristics. This is Item 14, which calls on students to evaluate the strengths and weaknesses of democratic systems by identifying the main message of a cartoon about differences of opinion in a democracy. The typical student across participating countries has a reasonable chance (about two in three) to respond correctly to tasks of this difficulty level (Appendix A, Table A.1).

The 75th percentile, separating the top quarter from the lower 75 percent in the distribution, was found at a scale value of 112. This difficulty level exceeds the difficulty parameters of all of the examples given above. Item 22, which asks about the functions of periodic elections and which also was used by IEA in 1971, comes closest, with a difficulty level of 113. Thus, a little less than one out of four students in the participating countries was—with the defined minimal degree of likelihood—able to respond correctly to this item or others of the same level of difficulty. In some countries, most notably in Poland, the national mean is almost at the level of the international 75th percentile.

The 95th percentile, finally, indicates the lower bound of achievement for the top 5 percent in the international distribution of civic knowledge. It is found at a scale value of 135, which implies that in this international top group the probability of success is higher than .65 for all the items in the test.³ This threshold is attained by the highest achieving students in a good

Panel 3.2 continued

number of participating countries (given here in the order of national means): the Slovak Republic, the Czech Republic, Australia, Hungary, Slovenia, Denmark, Germany, England, Sweden and Bulgaria. Another group of countries where the top 5 percent of the students were even more successful on the test includes Poland, Finland, Cyprus, Greece, Hong Kong (SAR), the United States, Italy, Norway and the Russian Federation.

The diagram in Figure 3.3 thus provides considerably more information than a simple ranking of countries by national average. The distributional properties for each country can and should be studied, preferably in close linkage with the appropriate statistics (such as the percentiles that were used here) and the item characteristics of the test.

Figure 3.4 provides a visual presentation of country averages showing those that differ significantly from each other at the 95 percent significance level. By selecting a country and reading across the table, we can see that the triangles pointing up indicate a significantly higher average performance than the country listed across the top, whereas the triangles pointing down stand for a significantly lower average. Dots indicate that the two country averages do not differ significantly from each other.⁴

The differences upon which this multiple-comparison table is based are mostly quite small between adjacent countries. For example, the country mean in Poland, where students have the highest average of all participating countries, does not significantly differ from the next seven countries. The Colombian average does not differ significantly from the Chilean one, but it does so from all other country means. Given that sampling errors are not identical across countries, it is possible that in some cases an apparently smaller difference is statistically significant while a larger one is not.

There is no obvious pattern in the distribution of national mean scores on the civic knowledge scale. Western European, North American and post-Communist countries lie in the top group of countries. The two Latin American countries and the Baltic countries as well as Romania have similar country means on the IEA civic knowledge test. We have already stated that many factors are likely to be involved in the emergence of between-country differences, even at the level of aggregate national indicators.

DISTINGUISHING BETWEEN CONTENT KNOWLEDGE AND INTERPRETATIVE SKILLS

As was shown above, the IEA test was designed to contain two different types of items: Type 1, tapping the students' knowledge of content; and Type 2, measuring their skills in interpreting civic-related material (cartoons, leaflets, descriptions of issues) and also incorporating their ability to distinguish between facts and opinions. It is also clear from the previous analysis that a one-dimensional representation of the students' response patterns is a psychometrically sound and meaningful way to present the findings. In this analysis each student is

Figure 3.4 Multiple Comparisons of Civic Knowledge

Instructions: Read *across* the row for a country to compare performance with the countries listed in the heading of the chart. The symbols indicate whether the mean achievement of the country in the row is significantly lower than that of the comparison country, significantly higher than that of the comparison country, or if there is no statistically significant difference between the two countries.

Country	Poland	Finland	Cyprus	Greece	Hong Kong (SAR)	United States	Italy	Slovak Republic	Norway	Czech Republic	Australia	Hungary	Slovenia	Denmark	Germany	Russian Federation	England	Sweden	Switzerland	Bulgaria	Portugal	Belgium (French)	Estonia	Lithuania	Romania	Latvia	Chile	Colombia
Poland		●	●	●	●	●	●	●	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
Finland	●		●	●	●	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
Cyprus	●	●		●	●	●	●	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
Greece	●	●	●		●	●	●	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
Hong Kong (SAR) ³	●	●	●	●		●	●	●	●	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
United States ¹	●	●	●	●	●		●	●	●	●	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
Italy	●	▼	●	●	●	●		●	●	●	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
Slovak Republic	●	▼	●	●	●	●		●	●	●	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
Norway ⁴	▼	▼	▼	▼	●	●	●		●	●	●	●	▲	▲	●	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
Czech Republic	▼	▼	▼	▼	●	●	●	●		●	●	●	●	●	●	●	▲	●	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
Australia	▼	▼	▼	▼	▼	▼	▼	▼	●	●		●	●	●	●	●	●	●	●	●	▲	▲	▲	▲	▲	▲	▲	▲
Hungary	▼	▼	▼	▼	▼	▼	▼	▼	●	●	●		●	●	●	●	●	●	▲	●	▲	▲	▲	▲	▲	▲	▲	▲
Slovenia	▼	▼	▼	▼	▼	▼	▼	▼	●	●	●	●		●	●	●	●	●	●	●	▲	▲	▲	▲	▲	▲	▲	▲
Denmark ⁴	▼	▼	▼	▼	▼	▼	▼	▼	●	●	●	●	●		●	●	●	●	●	●	▲	▲	▲	▲	▲	▲	▲	▲
Germany ²	▼	▼	▼	▼	▼	▼	▼	▼	●	●	●	●	●	●		●	●	●	●	●	▲	▲	▲	▲	▲	▲	▲	▲
Russian Federation ³	▼	▼	▼	▼	▼	▼	▼	▼	●	●	●	●	●	●	●		●	●	●	●	●	●	●	●	●	●	●	●
England ¹	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	●	●	●	●	●	●		●	●	●	▲	▲	▲	▲	▲	▲	▲	▲
Sweden ¹	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	●	●	●	●	●	●	●		●	●	●	▲	▲	▲	▲	▲	▲	▲
Switzerland	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	●	▼	●	●	●	●	●	●		●	●	●	▲	▲	▲	▲	▲	▲
Bulgaria	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	●	●	●	●	●	●	●	●	●		●	●	●	●	●	▲	▲	▲
Portugal ⁵	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	●	▼	●	●	●		●	●	●	▲	▲	▲	▲
Belgium (French) ⁴	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	●	▼	▼	●	●	●		●	●	●	●	▲	▲
Estonia	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	●	●	●		●	●	▲	▲
Lithuania	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	●	●	●	●		●	●	▲
Romania	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	●	●	●		●	▲
Latvia	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	●	●	●	●		▲
Chile	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	●		●
Colombia	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	●

Countries are ordered by mean achievement across the heading and down the rows.

- ▲ Mean achievement significantly higher than comparison country.
- No statistically significant difference from comparison country.
- ▼ Mean achievement significantly lower than comparison country.

NOTE: Significance tests at .05 level, adjusted for multiple comparisons.

- 1 Countries with testing date at beginning of school year.
- 2 National Desired Population does not cover all International Desired Population.
- 3 Countries did not meet age/grade specification.
- 4 Countries' overall participation rate after replacement less than 85 percent.
- 5 In Portugal, Grade 8 selected instead of Grade 9 due to average age. Mean scale score for Grade 9 was 106.

Source: IEA Civic Education Study, Standard Population of 14-year-olds tested in 1999.

simply awarded one total civic knowledge score, and from these scores the reported national means and distributions presented in the last section are derived.

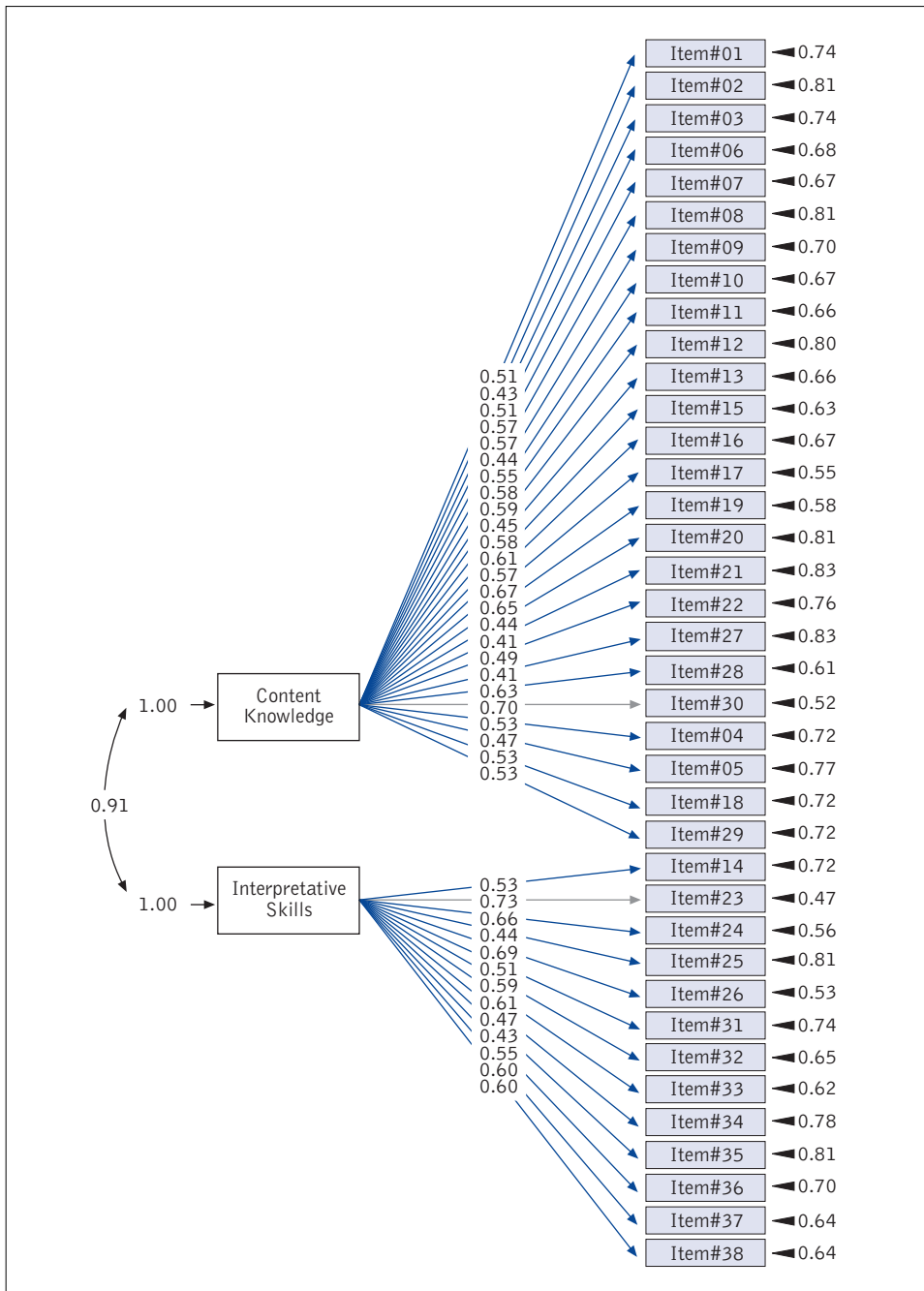
Even though this method is completely defensible, it was of interest to investigate whether the distinction between the two item types would map the students' response patterns even more appropriately. Might it be possible to derive a sub-score for content knowledge and for skills in interpreting civic-related material that would allow additional insights? To ascertain the dimensional structure of the IEA test under the assumption that two dimensions (corresponding to content knowledge and interpretative skills) could be distinguished, we performed confirmatory factor analyses (CFA). These were based on a calibration sample of 500 randomly selected students per country. As a contrast to the one-dimensional model with all 38 items on a single factor, we estimated a second model that allocated the 25 Type 1 items to one factor (content knowledge) and the remaining 13 items to a second factor (interpretative skills).⁵

Figure 3.5 shows a graphical display of the two-factor-structure and the results of a comparison of the model fit with the one-factor-solution. The diagram includes the factor correlation between the two latent constructs 'knowledge of content' and 'skills of interpretation', as well as the factor loading for each item and the proportion of (unexplained) error. The variance of the items explained by the latent dimensions ranges from 17 to 53 percent.

Both models show good statistical properties, but the two-factor-solution that is shown graphically has a relatively better model fit than the one-factor-solution. The difference in the chi-square statistic is 604, with one degree of freedom, which is highly significant.⁶ However, the two factors are strongly correlated with each other ($r = .91$), which indicates that the two abilities—'content knowledge' and 'interpretative skills'—refer to highly similar but not identical aspects of student performance. While it is true that the students who know much about civic-related content are likely also to have highly developed skills in interpreting civic-related material, this is not always the case. It is also possible that systematic differences occur between groups of students and even countries. Based on these results, we decided to present, in addition to the total cognitive score on *civic knowledge*, findings based on the two component sub-scales *content knowledge* and *skills in interpreting civic-related information*.

In Figure 3.6, we have again ranked the participating countries in the order of average achievement in civic knowledge (total score), but we also have given the means for the sub-scales of content knowledge and interpretative skills. All three scales are one-parameter Rasch scales, and all three have been set to an international mean of 100 and an international standard deviation of 20. It is not our assumption that one of the sub-scales refers to more complex or superior abilities than the other. On the contrary: in recognition of the fact that scale properties always depend on the particular choice of items used, our underlying assumption is that there is no meaningful way of comparing directly the scores on one of the sub-scales with those on the other. Given, however, that the students who took the items from the two sub-scales were identical, it is meaningful to compare the performance of countries on the two sub-scales, *relative to the respective international means*. In a sense, this analysis amounts to investigating the differential functioning of two groups of items that can be considered as measuring two different aspects of civic knowledge.

Figure 3.5 Confirmatory Factor Analysis of Civic Knowledge Items



Comparison of Model Fit for One- and Two-Factor-Solution*

Model	RMSEA	RMR	AGFI	CFI	NNFI	Chi-Square / DF
One-Factor-Solution	0.045	0.031	0.92	0.89	0.89	11329 / 665
Two-Factor-Solution	0.041	0.029	0.93	0.91	0.90	10725 / 664
Model Comparison	604 / 1					

* Models estimated with Weighted Least Square (WLS) and tetrachoric correlation matrices. Chi-square values corrected for non-normality. Calibration sample with 500 students per country.

Source: IEA Civic Education Study, Standard Population of 14-year-olds tested in 1999.

Figure 3.6 demonstrates that the outcomes on the two sub-scales do vary somewhat within countries. In 26 countries (the exceptions are Denmark and Germany), either the content or skills sub-score average is significantly above or below the international mean. In contrast, the total civic knowledge score reveals that 18 of the countries are above or below the international mean. This suggests that looking at two sub-scores rather than a single civic knowledge score does contribute to an understanding of relative strengths and weaknesses of civic knowledge as developed in the participating countries.

In five countries—Australia, England, Sweden, Switzerland and the United States—relative to the sub-scale means, the items pertaining to the skills dimension are more likely to be answered correctly than those from the content dimension. In two countries, Poland and the Russian Federation, relative performance on the content items is better than that on the skills items.

Once again these findings depend heavily on the specific items in the respective sub-scales and therefore can only be interpreted as pertaining to the *relative* position of countries on the two sub-scales. Nevertheless, it is interesting that, in some countries, students have acquired their content knowledge and interpretative skills differentially. In the absence of any obvious explanation, it is all the more important to systematically investigate such patterns in future analyses and to relate results to the findings from Phase 1 (Torney-Purta, Schwille & Amadeo, 1999; Steiner-Khamsi, Torney-Purta & Schwille, forthcoming).

GENDER DIFFERENCES IN CIVIC KNOWLEDGE

Most previous research has shown gender differences regarding cognitive performance of students in this area. In particular, according to studies conducted a decade or more ago, males scored higher on civic knowledge tests than females. In the first IEA Civic Education Study (1971), gender had a significant effect on cognitive achievement among the 14-year-old students in four out of eight countries, and these gender differences became more notable among older students, with males consistently outscoring females (Torney *et al.*, 1975, p.148). National studies mostly have had similar results on different tests measuring political knowledge (see Panel 3.1).

Figure 3.7 shows the gender differences in civic knowledge for all 28 countries in the present study. After correcting for multiple comparisons, we found no statistically significant differences in 27 of the 28 countries.⁷ The exception is Slovenia, where females perform better than males.

These findings suggest that, at least in a simple comparison, among 14-year-olds in most countries, political content knowledge and skills in interpreting political communication are unrelated to gender and that the previously found dominance of males in this area, even at the age of 14, is greatly diminished. This interpretation needs some modification when other variables are taken into account, however. More about this topic will be said in Chapter 8.

Figure 3.6 Content Knowledge Subscore and Interpretative Skills Subscore

Country	Mean Scale Scores			80 100 120
	Content Knowledge	Interpretative Skills	Total Civic Knowledge	
Poland	▲ 112 (1.3)	▲ 106 (1.7)	▲ 111 (1.7)	
Finland	▲ 108 (0.7)	▲ 110 (0.6)	▲ 109 (0.7)	
Cyprus	▲ 108 (0.5)	▲ 108 (0.5)	▲ 108 (0.5)	
Greece	▲ 109 (0.7)	▲ 105 (0.7)	▲ 108 (0.8)	
Hong Kong (SAR) ³	▲ 108 (1.0)	▲ 104 (1.0)	▲ 107 (1.1)	
United States ¹	● 102 (1.1)	▲ 114 (1.0)	▲ 106 (1.2)	
Italy	▲ 105 (0.8)	▲ 105 (0.7)	▲ 105 (0.8)	
Slovak Republic	▲ 107 (0.7)	▲ 103 (0.7)	▲ 105 (0.7)	
Norway ⁴	▲ 103 (0.5)	▲ 103 (0.4)	▲ 103 (0.5)	
Czech Republic	▲ 103 (0.8)	● 102 (0.8)	▲ 103 (0.8)	
Australia	● 99 (0.7)	▲ 107 (0.8)	● 102 (0.8)	
Hungary	▲ 102 (0.6)	● 101 (0.7)	● 102 (0.6)	
Slovenia	▲ 102 (0.5)	● 99 (0.4)	● 101 (0.5)	
Denmark ⁴	● 100 (0.5)	● 100 (0.5)	● 100 (0.5)	
Germany ²	● 99 (0.5)	● 101 (0.5)	● 100 (0.5)	
Russian Federation	● 102 (1.3)	▼ 96 (1.3)	● 100 (1.3)	
England ¹	▼ 96 (0.6)	▲ 105 (0.7)	● 99 (0.6)	
Sweden ¹	▼ 97 (0.8)	▲ 102 (0.7)	● 99 (0.8)	
Switzerland	▼ 96 (0.8)	● 102 (0.8)	● 98 (0.8)	
Bulgaria	● 99 (1.1)	▼ 95 (1.3)	● 98 (1.3)	
Portugal ⁵	▼ 97 (0.7)	▼ 95 (0.7)	▼ 96 (0.7)	
Belgium (French) ⁴	▼ 94 (0.9)	▼ 96 (0.9)	▼ 95 (0.9)	
Estonia	▼ 94 (0.5)	▼ 95 (0.5)	▼ 94 (0.5)	
Lithuania	▼ 94 (0.7)	▼ 93 (0.7)	▼ 94 (0.7)	
Romania	▼ 93 (1.0)	▼ 90 (0.7)	▼ 92 (0.9)	
Latvia	▼ 92 (0.9)	▼ 92 (0.8)	▼ 92 (0.9)	
Chile	▼ 89 (0.6)	▼ 88 (0.8)	▼ 88 (0.7)	
Colombia	▼ 89 (0.8)	▼ 84 (1.2)	▼ 86 (0.9)	

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

▲ Country mean significantly higher than international mean.

● No statistically significant difference between country mean and international mean.

▼ Country mean significantly lower than international mean.

◇ = Mean Subscore Civic Content (± 2 SE).

● = Mean Subscore Civic Skills (± 2 SE).

■ = Mean Total Civic Knowledge Score.

1 Countries with testing date at beginning of school year.

2 National Desired Population does not cover all International Desired Population.

3 Countries did not meet age/grade specification.

4 Countries' overall participation rate after replacement less than 85 percent.

5 In Portugal, Grade 8 selected instead of Grade 9 due to average age. Mean scores for Grade 9 were 108 on the Civic Content Scale, 103 on the Civic Skills Scale and 106 on the Total Civic Knowledge Scale.

Source: IEA Civic Education Study, Standard Population of 14-year-olds tested in 1999.

Figure 3.7 Gender Differences in Civic Knowledge

Country	Mean Scale Score Females	Mean Scale Score Males	Difference Absolute Value	Gender Difference		
				+10	0	+10
Denmark ⁴	99 (0.7)	102 (0.7)	3 (1.0)	Males Score Higher		Females Score Higher
Switzerland	97 (0.8)	100 (0.9)	2 (1.2)			
Chile	88 (0.8)	89 (0.8)	2 (1.1)			
Czech Republic	102 (0.8)	104 (1.0)	2 (1.3)			
Portugal ⁵	96 (0.8)	97 (0.9)	1 (1.2)			
Germany ²	99 (0.6)	101 (0.7)	1 (0.9)			
Norway ⁴	103 (0.6)	103 (0.7)	1 (0.9)			
Russian Federation ³	99 (1.2)	100 (1.7)	0 (2.1)			
Slovak Republic	105 (0.8)	105 (0.9)	0 (1.1)			
England ¹	99 (0.8)	100 (1.0)	0 (1.3)			
Cyprus	108 (0.7)	108 (0.6)	0 (0.9)			
Colombia	87 (1.3)	86 (1.1)	0 (1.7)			
Romania	92 (1.0)	91 (0.9)	0 (1.4)			
Hungary	102 (0.7)	101 (0.8)	1 (1.0)			
Hong Kong (SAR) ³	108 (1.1)	106 (1.4)	1 (1.8)			
Sweden ¹	100 (0.8)	99 (1.1)	1 (1.3)			
Estonia	95 (0.6)	93 (0.7)	1 (0.9)			
Finland	110 (0.9)	108 (0.8)	2 (1.2)			
United States ¹	107 (1.2)	106 (1.3)	2 (1.8)			
Greece	109 (0.8)	107 (0.9)	2 (1.2)			
Italy	106 (0.9)	104 (1.1)	2 (1.4)			
Bulgaria	99 (1.5)	97 (1.2)	2 (2.0)			
Lithuania	95 (0.8)	92 (0.8)	2 (1.1)			
Australia	103 (0.9)	101 (1.1)	2 (1.4)			
Poland	112 (2.2)	109 (1.5)	3 (2.6)			
Slovenia	102 (0.6)	99 (0.6)	4 (0.8)	Gender difference statistically significant at .05 level.		Gender difference not statistically significant.
Latvia	93 (0.9)	90 (0.9)	4 (1.3)			
Belgium (French) ⁴	97 (1.1)	93 (1.3)	5 (1.7)			

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

International Means		
Female	Male	Difference
100.4	99.7	0.7
(Averages of all country means)		

- Gender difference statistically significant at .05 level.
- Gender difference not statistically significant.

- 1 Countries with testing date at beginning of school year.
- 2 National Desired Population does not cover all International Desired Population.
- 3 Countries did not meet age/grade specification.
- 4 Countries' overall participation rate after replacement less than 85 percent.
- 5 In Portugal, Grade 8 selected instead of Grade 9 due to average age.

Source: IEA Civic Education Study, Standard Population of 14-year-olds tested in 1999.

HOME LITERACY RESOURCES AND CIVIC KNOWLEDGE

Previous research has consistently shown students' social background to be important in fostering civic knowledge. Students from less affluent and less educated families have less knowledge in this area than those with more affluent and better-educated parents. In the first IEA Civic Education Study, social status was a consistently positive predictor for the cognitive test score in all participating countries (Torney *et al.*, 1975).

For international studies it is very difficult to find comparable indicators for the social background of families. Social status is defined differently across countries, and it is usually impractical to measure social status by asking 14-year-old students about their parents' occupation(s) or family income. The Student Background Questionnaire of the present study included questions to the students as to their father's and mother's educational attainment, but the validity of this indicator may not be beyond question either. Students of this age sometimes do not know the educational level of their parents: in some of the participating countries more than 20 percent of the students did not answer this question. Another serious problem is that the different structure of educational systems across countries may jeopardize the comparability of the educational levels.

We therefore decided to use the number of books in the home as the indicator for the students' social background. This variable has been used before in international studies on educational achievement and has proven to be a very consistent predictor of educational achievement (see, for example, the reports of TIMSS, Beaton *et al.*, 1996). The number of books in the home can be interpreted as a proxy for the emphasis placed on education, the resources available to acquire and support literacy and, more generally speaking, the academic support a student finds in his or her family.

Table 3.1 shows that home literary resources are quite consistently correlated with the civic knowledge score. The inspection of squared E_{tas} —a measure of the proportion of variance in the (dependent) knowledge scores attributable to the (independent) variable 'number of books in the home'—reveals that in all but one country (Hong Kong/SAR) home literacy resources account for more than 5 percent of the variance in the test scores. The strongest effects of 12 percent and more are found in Chile, England, Germany, Hungary and the United States. In the large majority of countries, the more books that students report in their homes, the better they perform on the civic knowledge test.

It should be noted that there are substantial differences in the students' report of home literacy resources across countries. Whereas in most European countries, Australia and the United States less than 10 percent of students report no or very few books in the home, in Chile, Colombia and Hong Kong (SAR) this is true of almost one-third. In Portugal and Romania, one-fifth of the students fall into this category. In Chile, Colombia and Romania this finding coincides with the fact that the overall average test scores are considerably below the international mean. As such, the low level of home literacy resources in these countries may be one reason among others for the relatively low performance on the test, at least as compared to other countries in this study. It may be recalled that a similar effect was suggested by the statistical relationship between national literacy rates and student average performance in the civic knowledge test.

Table 3.1 Civic Knowledge and Students' Reports on Home Literacy Resources

Country	None or Very Few (0-10 books)		About One Shelf (11-50 books)		About One Bookcase (51-100 books)		About Two Bookcases (101-200 books)		Three or More Bookcases (More than 200 books)		Eta Squared
	Percent of Students	Mean Scale Score	Percent of Students	Mean Scale Score	Percent of Students	Mean Scale Score	Percent of Students	Mean Scale Score	Percent of Students	Mean Scale Score	
Australia	4 (0.4)	88 (1.4)	14 (0.8)	95 (1.2)	20 (0.9)	100 (1.0)	23 (1.0)	104 (1.1)	39 (1.3)	106 (1.0)	0.05
Belgium (French)	10 (1.1)	83 (1.1)	18 (1.1)	90 (1.2)	22 (0.9)	93 (1.1)	18 (0.9)	97 (1.0)	32 (1.5)	102 (1.6)	0.11
Bulgaria	14 (1.8)	86 (1.5)	15 (1.4)	92 (1.4)	20 (1.6)	96 (1.3)	19 (0.8)	100 (1.7)	32 (1.8)	105 (1.9)	0.11
Chile	32 (1.6)	80 (0.6)	33 (1.0)	89 (0.7)	19 (0.7)	94 (0.8)	8 (0.6)	97 (0.9)	8 (0.5)	96 (1.1)	0.13
Colombia	29 (2.5)	81 (1.1)	33 (1.2)	87 (0.9)	21 (1.4)	90 (1.3)	10 (0.9)	92 (1.3)	7 (0.7)	90 (1.5)	0.06
Cyprus	8 (0.5)	97 (1.2)	27 (0.8)	105 (0.7)	31 (0.9)	109 (0.8)	18 (0.6)	112 (0.9)	15 (0.8)	113 (1.1)	0.06
Czech Republic	2 (0.3)	93 (2.5)	12 (0.8)	93 (0.9)	29 (1.1)	99 (0.9)	31 (1.3)	105 (0.8)	27 (1.3)	110 (1.2)	0.07
Denmark	6 (0.5)	89 (1.3)	17 (0.8)	95 (1.0)	23 (0.8)	98 (0.8)	21 (0.7)	103 (0.9)	32 (0.9)	106 (0.7)	0.06
England	8 (0.7)	86 (0.9)	19 (0.8)	92 (0.7)	23 (0.9)	97 (0.9)	22 (0.9)	102 (0.8)	27 (1.3)	109 (0.9)	0.15
Estonia	1 (0.2)	83 (2.2)	7 (0.5)	89 (1.2)	17 (0.8)	89 (0.6)	27 (0.9)	93 (0.6)	49 (1.2)	98 (0.8)	0.05
Finland	5 (0.5)	96 (1.5)	24 (0.9)	105 (0.8)	30 (0.9)	109 (0.8)	22 (0.8)	112 (0.8)	20 (1.1)	116 (1.4)	0.05
Germany	6 (0.5)	86 (1.3)	21 (0.7)	93 (0.7)	24 (0.8)	97 (0.6)	19 (0.7)	103 (0.7)	30 (1.0)	108 (0.9)	0.14
Greece	8 (0.7)	95 (1.4)	34 (1.0)	104 (0.8)	27 (0.8)	109 (0.9)	15 (0.8)	113 (1.4)	15 (0.8)	116 (1.2)	0.07
Hong Kong (SAR)	31 (1.0)	103 (1.2)	35 (0.9)	111 (1.2)	16 (0.6)	107 (1.4)	8 (0.4)	109 (1.7)	9 (0.5)	104 (1.6)	0.02
Hungary	5 (0.6)	86 (1.2)	12 (0.8)	91 (0.7)	21 (1.0)	97 (0.8)	24 (0.9)	103 (0.7)	38 (1.4)	108 (0.8)	0.14
Italy	15 (0.8)	95 (0.9)	30 (1.1)	103 (0.8)	24 (0.7)	107 (0.9)	16 (0.8)	111 (1.1)	14 (0.9)	113 (1.5)	0.08
Latvia	2 (0.4)	82 (1.7)	10 (0.8)	86 (1.2)	20 (0.9)	88 (1.0)	26 (1.1)	91 (1.0)	40 (1.6)	96 (1.1)	0.05
Lithuania	8 (0.7)	85 (1.6)	23 (1.0)	90 (1.0)	26 (0.9)	93 (0.7)	18 (0.7)	96 (0.9)	24 (1.1)	100 (1.0)	0.07
Norway	3 (0.4)	88 (1.5)	15 (0.8)	96 (0.9)	21 (0.9)	100 (0.7)	26 (0.8)	104 (0.7)	35 (1.3)	109 (0.7)	0.07
Poland	6 (0.5)	94 (1.3)	20 (1.2)	103 (1.6)	24 (1.3)	109 (2.1)	19 (1.1)	113 (1.8)	31 (2.2)	119 (1.9)	0.11
Portugal	22 (1.3)	90 (0.6)	36 (1.2)	94 (0.6)	20 (0.8)	98 (1.0)	11 (0.8)	103 (1.1)	11 (1.1)	106 (2.1)	0.09
Romania	24 (1.9)	87 (1.7)	28 (1.2)	90 (0.9)	21 (1.0)	91 (0.9)	12 (0.9)	96 (1.1)	14 (1.1)	98 (1.5)	0.05
Russian Federation	6 (1.0)	84 (3.2)	17 (1.4)	94 (1.6)	26 (1.0)	98 (1.6)	21 (1.0)	102 (1.6)	30 (1.4)	106 (1.7)	0.08
Slovak Republic	4 (0.5)	90 (1.6)	21 (1.0)	100 (1.0)	30 (1.1)	104 (0.7)	24 (1.0)	108 (1.0)	21 (1.3)	113 (1.1)	0.09
Slovenia	7 (0.8)	89 (0.8)	29 (1.0)	96 (0.6)	30 (0.9)	101 (0.7)	17 (0.9)	105 (1.0)	17 (0.8)	107 (1.0)	0.08
Sweden	5 (0.5)	84 (1.3)	16 (0.9)	93 (1.2)	24 (1.2)	98 (1.1)	23 (1.2)	99 (0.9)	32 (1.9)	106 (1.1)	0.09
Switzerland	6 (0.7)	86 (1.0)	24 (1.3)	93 (0.9)	25 (0.9)	98 (0.9)	21 (0.8)	102 (1.0)	23 (1.2)	105 (1.3)	0.10
United States	9 (0.9)	91 (1.3)	22 (1.2)	99 (0.9)	22 (0.8)	105 (1.3)	20 (1.0)	112 (1.5)	28 (1.4)	115 (1.7)	0.12
International Sample	10 (0.2)	89 (0.3)	22 (0.2)	96 (0.2)	23 (0.2)	100 (0.2)	19 (0.2)	103 (0.2)	25 (0.2)	106 (0.3)	0.07

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

Source: IEA Civic Education Study, Standard Population of 14-year-olds tested in 1999.

Hong Kong (SAR), however, shows a very different pattern. Civic knowledge does not increase significantly with home literacy resources: most students report a low number of books at home and the overall test score is above the international average, a pattern which is parallel to findings from TIMSS (Beaton *et al.*, 1996). Conversely, students from some countries, for example, Belgium (French) and the Baltic countries, report a relatively high number of books in their homes, but their average performance on the civic knowledge test is below the international mean. These exceptions demonstrate once more the need for more specific investigation that goes beyond the general trends appearing in this first cross-national volume.

SUMMARY

The analysis of responses to the IEA test of civic knowledge presented in this chapter shows that students do vary in their civic knowledge (and in sub-scores distinguishing content knowledge and interpretative skills). Differences between countries are smaller than in such school subjects as mathematics where the acquisition of the respective abilities is likely to depend more on the instruction received, and are relatively similar to studies in reading literacy. Most of the students in the participating countries give evidence of a fairly adequate content knowledge base when questioned about basic notions of democracy and citizenship. However, the analysis also reveals that in some countries there is a considerable group of students whose civic knowledge is below the level that might be considered 'basic'. At the same time, it should be noted that the questions in the test answered correctly by only relatively small numbers of students are part of what might be required to perform such civic tasks as deciding between candidates based on their election leaflets, understanding newspaper editorials and deciding whether to join a political organization with a particular ideology.

Gender differences in civic knowledge that were prominent in earlier research are not manifest in the present data. As more recent studies have also shown, boys no longer seem to have a clear advantage over girls in their knowledge of political content and their skills in interpreting civic-related materials. The question will be addressed again in Chapter 8, which considers the potentially moderating effects of other variables.

Knowledge differences between students that can be traced back to differences in the learning environments of the homes persist, however. In other international achievement surveys, home literacy resources have been closely related to the test results, and here, too, it remains to be seen how much of this effect stands if it is investigated in conjunction with other factors related to the emergence of civic knowledge among 14-year-olds.

NOTES

- 1 The relationship between the item scale value (difficulty parameter) and the percentage of correct answers is not as direct as it seems here. Thus, the idea that a student with a scale score of 106 has a .65 percent chance of scoring correctly on item example 3 appears to conflict with the fact that the United States' mean scale score is 106, with the percentage correct being 53, not 65. Likewise, on item example 5, Australia's mean scale score is 102, with a percentage correct of 75, not 65. Reasons for such discrepancies can be found in distributional properties, minor item-by-country interactions, and the fact that parameter estimates are based on the entire response matrix that includes local deviations from the model assumptions.
- 2 The adjustment procedure renders a gross measure of annual growth in civic knowledge that is in the order of 40 percent of an international standard deviation and thus roughly in line with the difference found between the two Portuguese samples (see Figure 3.3, footnote 5).
- 3 This observation is compatible with the fact that the most difficult item in the test has a scale value of only 121. If we take the case of a student who answered all the items correctly, we can see that his or her ability level is 'unknown' (and therefore set to some pragmatically plausible value), given that there were no items difficult enough to facilitate an ability estimate at the defined probability level of .65. Similarly, and with good mathematical reason, students with near-perfect test results can receive scores well above the item difficulty parameter of the most difficult item.
- 4 All significance tests ($p < .05$) are based on the Dunn-Bonferroni procedure for multiple comparisons, applied here to one country at a time ($df = 27$). Differences in rounding conventions used in different statistical packages may result in small inconsistencies.
- 5 Because all items considered are dichotomously keyed (correct vs. incorrect response), we chose a conservative approach. We estimated the two models according to the 'weighted least squares estimation' (WLE) algorithm, using a matrix of tetrachoric correlations and an asymptotic covariance weight matrix (see Jöreskog, 1990).
- 6 However, in view of the large sample size of 14,000 students (500 randomly selected students from each participating country), the interpretation of the chi-square statistic is generally problematic because even small differences tend to be significant.
- 7 Comparisons within a single country that are not corrected in this way may show a significant difference.