

[2012]

**Report of the Data Analysis of the Student  
Questionnaire Used to Measure Teacher Factors :  
Across and Within Country Results**



**Leonidas Kyriakides  
Bert Creemers  
Anastasia Panayiotou**

## **ACKNOWLEDGEMENTS**

The analysis procedures and results presented in this report are part of a 3-year project (2009-2012) entitled “Establishing a knowledge-base for quality in education: Testing a dynamic theory of educational effectiveness” (08-ECRP-012), funded by the Cyprus Research Promotion Foundation and the European Science Foundation.

# Table of Content

<b>1.</b>	<b>Across country results</b>	
<b>1.1</b>	<b>Steps of data processing for the analysis of the student questionnaire data.....</b>	<b>4</b>
1.1.1	Cleaning the data.....	4
1.1.2	Recoding the data.....	5
1.1.3	Conducting Reliability Analysis.....	5
<b>1.2</b>	<b>Generalisability Analysis.....</b>	<b>5</b>
<b>1.3</b>	<b>Categorization of items according to factors and dimensions.....</b>	<b>6</b>
<b>1.4</b>	<b>Confirmatory Factor Analysis (CFA).....</b>	<b>7</b>
<b>1.5</b>	<b>Searching for Grouping of Factors: A Model Describing Quality of Teaching.....</b>	<b>10</b>
<b>2.</b>	<b>Within country results</b>	
<b>2.1</b>	<b>Generalisability Analysis.....</b>	<b>16</b>
<b>2.2</b>	<b>Confirmatory Factor Analysis.....</b>	<b>16</b>
<b>2.3</b>	<b>Within Countries SEM analysis.....</b>	<b>23</b>
	<b>References.....</b>	<b>30</b>
	<b>Appendix A.....</b>	<b>31</b>
	<b>Appendix B.....</b>	<b>42</b>
	<b>Appendix C.....</b>	<b>52</b>
	<b>Appendix D.....</b>	<b>53</b>

# **1. Across country results**

The first two sections of the report present the results from the across-countries and the within country analyses, correspondingly.

## **1.1 Steps of data processing for the analysis of the student questionnaire data**

In this part of the data documentation steps of data cleaning and preparation are described. Specifically three steps were undertaken for this task: (1) cleaning the data delivered by all the countries, (2) recoding the data, (3) conducting Reliability analysis for the whole scale. These data cleaning and analysis procedures are described in detail below.

### **1.1.1 Cleaning the data**

As part of the first steps of data cleaning, descriptive statistics by item were conducted in order to check carefully whether any mistakes were made regarding the coding of the questionnaire data. According to the coding guidelines that were given to all the countries the coding for all items was from 1 to 5 (1= “Never” – 5= “Almost Always”). Missing values were indicated by using the codes 7, 8, and 9: The code 9 was given when a student omitted the task. The code 8 was used when a student’s response was ambiguous. The code 7 was used to indicate that items were not administered. Where a mismatch of data coding was found the corresponding country was notified and the data were being corrected. The number of missing values per item is presented in Appendix A, Table A1. The percentages of the items that were coded with 7, 8 and 9 were very low therefore they were considered as missing and no additional processes were made. More specifically, as can be seen in Table A1 (see Appendix A), for code 7 there were only 139 missing values for each item and all of them were located in Belgium (see table A2) since the student questionnaire was not administered to three classrooms of the sample. For code 8 the number of missing values was very small (the largest number of missing values with code 8 was 26 for item 26) and for code 9 the percentage of missing values was not more than 8%. The only exception, with a larger percent of missing values with code 9 was for item 25 as that item was not administered at all in Germany (see table A4). Therefore, no

approach for replacing the missing values was used. Obviously, item 25 was excluded from the analyses in Germany.

### **1.1.2 Recoding the data**

As mentioned, all the items were coded by all the participating countries from 1 to 5. However, some of the items had a negative meaning (items 12, 23, 31, 36, 25, 39, 41, 43, 18, 27, 30, 29, 33). Negatively worded items are those phrased in the opposite semantic direction from the majority of the items on a measure (Barnette, 2000). These negative items were included in the questionnaire since as Barnette (2000) mentions *“negatively worded and correspondingly reverse-scored items have been used extensively in survey constitution to guard against acquiescent behaviors or the tendency for respondents to generally agree with survey statements more than disagree”*. Therefore, these items had to be recoded in order to match the answers of the rest of the scale.

### **1.1.3 Conducting Reliability Analysis**

After recoding the negatively worded items reliability analysis was conducted and the Cronbach alpha was calculated for the entire scale (49 items included in part A of the questionnaire). The results of the reliability analysis showed that the Cronbach alpha was satisfactory ( $\alpha = 0.786$ ). In addition, the calculation of the value of the Alpha “if item deleted” revealed that none of the items had to be removed.

## **Generalisability Analysis**

After the cleaning and preparation of the data, a Generalisability Study on the use of students’ ratings was conducted (Shavelson, Webb & Rowley, 1989; Cronbach, Gleser, Nanda, & Rajaratnam, 1972). The results of the ANOVA analysis (see Appendix A, Table A8) showed that the data can be generalized at the classroom level, as for all the items of the questionnaire, the between group variance was higher than the within group variance ( $p < 0.05$ ).

## **1.2 Categorization of items according to factors and dimensions**

The questionnaire was translated and adapted from an earlier version that was used in previous studies in Cyprus and in which all the factors and dimensions were measured (for the questionnaire see Appendix B). The validity of the original questionnaire with the five dimensions was tested and the results were satisfactory (Kyriakides & Creemers, 2008, Creemers & Kyriakides, 2010). However, in spite of the fact that in the original questionnaire that was used in earlier studies in Cyprus all the dimensions were measured, for this project some of the questionnaire items had to be removed for adaptation reasons as they did not match the context of some participating countries. Therefore, it was not possible to conduct factor analyses to measure the importance of each dimension (for the specification table with the categorization of items in the five dimensions, see Appendix C). Therefore, in Table 2 we have grouped the items into those concerned with dimensions measuring quantitative (i.e., frequency and stage) and qualitative characteristics of each factor (i.e., focus, quality, differentiation). Our intention was to test each of the teacher factors with respect to these two broader categories.

**Table 2: Quantitative and Qualitative Characteristics of the Teacher Factors**

	Dimensions	
Teacher Factors	Quantitative Characteristics (Frequency, Stage)	Qualitative Characteristics (Focus, Quality, Differentiation)
Structuring	3, 2, 34, 38	1, 4, 7, 10
Application	11, 12	26, 13, 14, 15, 32
Management of Time	31, 35, 36	Not applicable (N/A)*
Questioning	25, 39	24, 37, 40, 41, 42, 43
Modeling	44, 47	45, 46
Classroom as a learning Environment / Teacher – Student Interaction	16, 17	19, 20, 21, 22
Classroom as a learning Environment / Dealing with Misbehaviour	29, 18	23, 27, 33, 30, 28
Assessment	50, 51	5, 6, 9, 48, 49

\***Note:** The factor concerned with the management of time is only measured by taking into account its frequency dimension (see Creemers & Kyriakides, 2008).

### 1.3 Confirmatory Factor Analysis (CFA)

Having in mind the categorization of the items in the Specification Table (see table 2), Confirmatory Factor Analysis was conducted for each of the teacher factors of the dynamic model which were measured by the student questionnaire by using the EQS software for Structural Equation Modeling (Bryne, 1994). CFA was used, as the objective was to test whether the data fit a hypothesized measurement model; in this case the assumptions of the dynamic model in regard to the two broader dimensions of each teacher factor. However, in the case of management of time, CFA was not conducted as there were only 3 items measuring the frequency dimension and the one-factor model is just identified (i.e., its degrees of freedom are 0). Therefore, for

management of time exploratory factor analysis was conducted with satisfactory results. Specifically, the first eigenvalue was equal to 1.40 and explained almost 50% of the variance whereas the second eigenvalue was much smaller than 1 (i.e., 0.81). These results show that we can treat these three items as belonging to one factor especially since all three items had relatively big loadings (i.e., bigger than 0.67). The CFA models which were conducted for the other factors, showed that some of the items that were included in the questionnaire had to be removed from the analyses. Therefore, the items that remained in each of the teacher factors in the CFA models and their dimensions are presented in Table 3. For the items that were excluded from the analyses some possible explanations are provided in section 1.4.

The reliability of each scale measuring these factors was also calculated and the results show that for each factor the Cronbach alpha was relatively satisfactory (given the small number of items of each subscale, see Cronbach, 1990). The only exception was the factor of questioning where Chronbach's alpha was small. In addition, the calculation of the value of the "Alpha if item deleted" revealed that none of the items had to be removed from each factor. The results of the reliability analysis per factor, across countries are also presented in Table 3. In addition, the covariance matrixes used for the SEM analysis were produced and SEM analysis per factor was initially conducted to find out whether the questionnaire items could help us develop scores for each factor. For the SEM analysis the EQS program was used. The fit indices of the one and two factor models that were produced are also presented in Table 3. In some factors (e.g., Modeling) a two-factor model could not be produced as there are only four items in the questionnaire measuring this factor but the single factor model was found to fit well to the data. Moreover, the items of the two aspects of the classroom as a learning environment factor were treated as belonging to two different one-factor models measuring the type of interactions that exist in the classroom and the teacher ability to deal with misbehaviour in a way that teaching time is not reduced.



**Table 3: Items of the CFA models and across countries results of the SEM and Reliability Analysis**

<b>Teacher Factors</b>	<b>Quantitative Characteristics</b>	<b>Qualitative Characteristics</b>	<b>Results: SEM and Reliability Analysis</b>
<b>Modeling</b>	44, 47	45, 46	One – Factor Model: ( $X^2= 47$ , $df= 2$ <b>CFI=0.991</b> , <b>RMSEA= 0.049</b> ) <b>a = 0.67</b>
<b>Structuring</b>	3, 2, 34, 38	1, 4, 7, 10	Two – Factor Model: ( $X^2= 298$ , $df= 18$ <b>CFI=0.967</b> , <b>RMSEA= 0.040</b> ) <b>a1 =0.62</b> , <b>a2= 0.52</b>
<b>Application</b>	11, 12	13, 26	One – Factor Model: ( $X^2= 2$ , $df= 1$ <b>CFI=0.99</b> , <b>RMSEA= 0.012</b> ) <b>a = 0.58</b>
<b>Questioning</b>	25, 39	24, 37, 40, 42	One – Factor Model: ( $X^2= 166$ , $df= 8$ <b>CFI=0.970</b> , <b>RMSEA= 0.045</b> ) <b>a = 0.46</b>
<b>Classroom as a learning Environment / Teacher – Student Interaction</b>	16, 17	19, 20, 21	One – Factor Model: ( $X^2= 21$ , $df= 5$ <b>CFI=0.997</b> , <b>RMSEA= 0.018</b> ) <b>a = 0.63</b>
<b>Classroom as a learning Environment / Dealing with Misbehaviour</b>	29	27, 33 , 30	One – Factor Model: ( $X^2= 230$ , $df= 2$ <b>CFI=0.965</b> , <b>RMSEA= 0.10</b> ) <b>a = 0.69</b>

**Note:** Some items had to be removed and these items are only presented in Table 2.

## **1.4 Searching for Grouping of Factors: A Model Describing Quality of Teaching**

Since one of the main assumptions of the dynamic model is that the factors are interrelated, the next step was to see how these effectiveness factors are related with each other. Specifically, our assumption was that the factors of management of time, teacher ability to deal with student misbehaviour and questioning: raising non-appropriate questions belong to one second order factor, whereas the other factors can be grouped to another second order factor. Therefore this assumption was tested by conducting SEM analysis to see whether the factors can be grouped according to our assumptions. Our attempt was to develop a model based on the data from all the countries and then to replicate this model in the within country analysis. From the SEM analysis, our assumption was confirmed and two second order factors were identified. The first second order factor consists of the factors measuring management of time, teacher ability to deal with student misbehaviour and questioning: raising non-appropriate questions. This overarching factor can be treated as the factor measuring the ability of teacher to maximize the use of teaching time and can be seen as an indicator of the quantity of teaching. All the other factors were found to load on the second overarching factor which can be treated as an indicator of the quality of teaching. It is finally important to note that the correlation between these two overarching factors is very small implying that those teachers who are able to use the teaching time are not necessarily able to use the teaching time effectively. Separate within country SEM analyses showed that, the model fits well to the data emerged by each country separately. The fit indices of the across country model are shown in Table 4.

Also, two other models were tested to compare their fit to the data with the proposed model. In the first model (Model 2) all the items that were used for the SEM analysis were considered as belonging to a single factor. This model was an attempt to see if the questionnaire items refer to a social desirability factor and thereby the questionnaire is not valid at all. In the second model (Model 3) the items concerning quality of teaching were considered as one factor and the items concerning quantity of teaching as another factor. If model 3 was found to fit to the data, this might cause doubts on whether we could have scores per each factor separately. The fit indices of

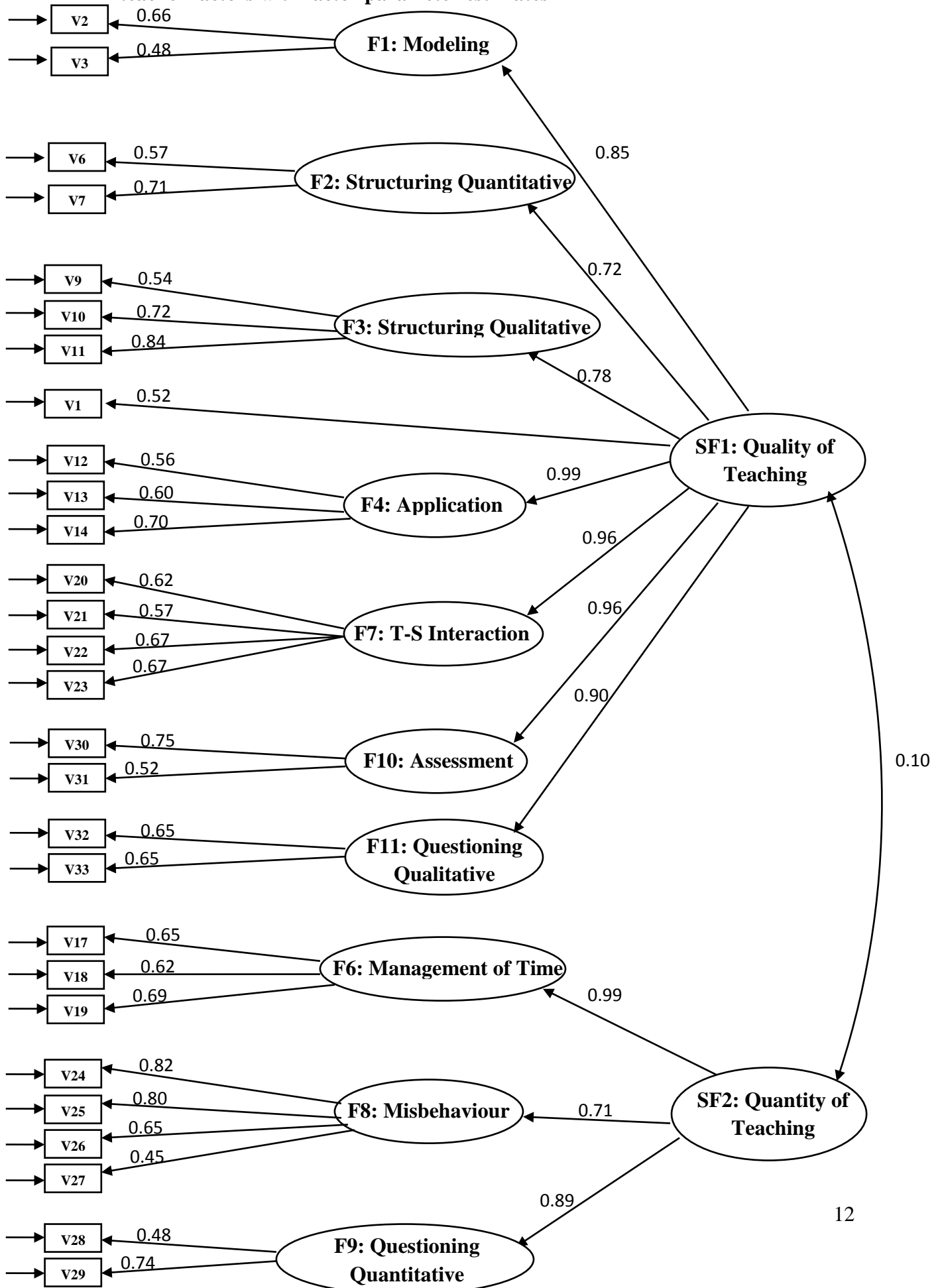
each model are shown in Table 4. We can see that model 1 was found to best fit the data and that the fit indices of model 1 were very good.

**Table 4: Results of the SEM analysis across countries**

<b>SEM analyses – Results</b>							
<b>Models for across countries (N= 9967)</b>	<b>X<sup>2</sup></b>	<b>Df</b>	<b>X<sup>2</sup>/df</b>	<b>p</b>	<b>CFI</b>	<b>RMSEA</b>	<b>Range RMSEA</b>
<b>Model 1</b>	3604	325	11,1	0.001	0.929	0.032	0.031 – 0.033
<b>Model 2</b>	16507	350	47,1	0.001	0.648	0.068	0.067 – 0.069
<b>Model 3</b>	6502	349	18,3	0.001	0.866	0.042	0.041 – 0.042

Based on the results of model 1 the factor scores were estimated based on the loadings of the items that occurred from the SEM analysis, as they appear below in Figure 1. These factor scores will be used for the multilevel analysis in order to identify their impact on student achievement in mathematics and science. Looking at the loadings of the items and the factors we can see that they are all very high and that all the loadings are statistically significant. Also, as it was noted earlier the correlation between the two overarching factors is very small which suggests that effectiveness in one overarching factor does not also imply effectiveness in the other.

**Figure 1: The second-order factor model of the student questionnaire measuring teacher factors with factor parameter estimates**



**Figure 1** presents the results from the across countries SEM analysis and shows the second-order factor model that fits the data of the student questionnaire. Below you can see explanations for the first and second order factors that are shown in the diagram and specifically:

**First Order Factors:**

**F1:** Modeling

**F2:** Structuring – Quantitative Characteristics

**F3:** Structuring – Qualitative Characteristics

**F4:** Application

**F6:** Management of Time

**F7:** Classroom as a Learning Environment – Qualitative characteristics: Teacher - Student interaction

**F8:** Classroom as a Learning Environment – Quantitative Characteristics: Dealing with Misbehaviour

**F9:** Questioning – Quantitative Characteristics: Raising non-appropriate questions

**F10:** Assessment

**F11:** Questioning – Qualitative Characteristics

**V1:** Orientation

**Second Order Factors:**

**SF1:** Quality of Teaching

**SF2:** Quantity of Teaching (Management of Time, Misbehaviour & Questioning: raising non – appropriate questions)

## **Interpretation of Results**

First, it is important to note that there was a difficulty in covering all five dimensions of the dynamic model, due to the fact that some of the items that were included in the original questionnaire had to be removed in order for the questionnaire to be suitable for the context of all the participating countries. Therefore, we had to collapse the five dimensions and create two which measure quantitative and qualitative characteristics of the functioning of each factor.

We have managed to create an overall model which shows the relations of the factors in all the countries but in order for it to best fit the data both within and across the countries, it had to be simpler and some items had to be removed.

From the CFA analyses (section 1.3), 15 items in total had to be removed. These items belong to four categories which are: a) negative items, b) items that include ***“I”***, c) items measuring the differentiation dimension and d) items with ***“when”***. From the items that were removed, 5 items were negative, 6 items included ***“I”***, 4 items were measuring the differentiation dimension and 8 items included the word ***“when”***. In total, the number of the removed items was higher than 15. This is because some items belonged to more than one category (i.e. When I finish a task before my classmates my teacher immediately gives me something else to do). However, all the removed items belonged to at least one category.

All the items that had to be removed from the analyses are presented in more detail below and they are classified based on our assumptions about the reasons for which they might have to be removed.

As mentioned, the results from the analyses showed that some of the negatively worded items had to be removed. A possible explanation could be that the students may not have understood them in order to answer them correctly or that translation may have caused some problems. These negative items were: 6, 23, 18, 41 and 43. In addition the items that included ***“I”*** (know) and had to be removed were items 28 and 7. This might be showing a problem in the generalisability of the data from these items as they are asking for the students’ opinion for their teacher, which can be problematic. Moreover, as it resulted from the CFA analyses some of the items

measuring the differentiation dimension had to be removed. This can probably be explained by the fact that differentiation dimension is not used in all the countries the same way and also by the fact that it is difficult for students to locate differentiation. The items measuring the differentiation dimension were: 13, 14, 15 and 32.

Additionally, items with “*when*” (e.g. When we are having a test) had to be removed, which shows that perhaps more specific options should have been given in each country so that the students would know which option to choose when the answer to these items was “not applicable”. These items were: 21, 22, 48 and 49.

At the next step we had to search for grouping of factors (see section 1.4) and our aim was to create a model that was comparable from country to country. In this attempt, 6 items had to be removed from 3 factors and they are presented below:

For Questioning factor: items 40 and 37 were removed from the SEM analysis and not from the factor analysis as they were found to behave differently in some countries.

For Modeling factor: items 46 and 47 were also removed from the SEM analysis as there loadings were found to be low and therefore their contribution to the factor was small.

And finally, for Structuring factor: items 2 and 38 were removed from the SEM analysis as there loadings were also found to be low.

In the next section, the within country analyses results are presented; it was again found that the model shown in Figure 1 is the one that best fits the data of each country separately.

## **2. Within country results**

A similar procedure as the across country analyses was used in order to conduct within country analyses where the data of each country were processed separately.

### **2.1 Generalisability Analysis**

A Generalisability Study on the use of students' ratings was conducted again with the data of each country separately. The results of the ANOVA analysis for each country showed that the data can be generalized at the classroom level as for all the items of the questionnaire, the between group variance was higher than the within group variance ( $p < 0.05$ ).

### **2.2 Confirmatory Factor Analysis (CFA)**

Then, Confirmatory Factor Analysis was conducted for each of the teacher factors of the dynamic model which were measured by the student questionnaire and their reliability was measured for each country. Again, no analyses were conducted for the orientation factor which was measured by only one item of the questionnaire. Also, as with the across countries analyses, CFA was not conducted for management of time for which there were only 3 items and the model was just identified ( $df=0$ ). Therefore, for management of time exploratory factor analysis was conducted by country with satisfactory results, which again showed that we can treat the items as one factor, as the loadings of the three items were very big (bigger than 0.60).

The CFA analysis by country which was conducted for the other factors produced similar results as the across country analysis; it showed that the same items included in the across country analysis should be included in the by country analysis for each of the teacher factors measured by the student questionnaire.

The reliability of these factors was also calculated for the data of each country separately and the Cronbach alpha of each factor was observed. The results of the reliability analysis showed that the Cronbach alpha was satisfactory for all factors. The only exception was in Belgium where Chronbach's alpha was relatively small for the factor concerned with the qualitative characteristics of structuring. In addition, the calculation of the value of the Alpha if item deleted" revealed that none of the items had to be removed from each factor. The results of the reliability analysis per factor,



within countries are presented in Tables 5-10. In addition, SEM analysis per factor was again conducted for each country and the fit indices of the one and two factor models that were produced from the within country analysis are also presented in Tables 5-10. In Germany item 25 was not administered and therefore was excluded from the country analyses.

**Table 5: Items of the CFA models and results of the SEM and Reliability Analysis for Belgium**

<b>Teacher Factors</b>	<b>Quantitative Characteristics</b>	<b>Qualitative Characteristics</b>	<b>Results: SEM and Reliability Analysis</b>
<b>Modeling</b>	44, 47	45, 46	One – Factor Model: ( $X^2= 168$ , $df= 2$ <b>CFI=0.944</b> , <b>RMSEA= 0.09</b> ) <b>a = 0.55</b>
<b>Structuring</b>	3, 2, 34, 38	1, 4, 7, 10	Two – Factor Model: ( $X^2= 46$ , $df= 19$ <b>CFI=0.978</b> , <b>RMSEA= 0.027</b> ) <b>a1 =0.63, a2= 0.36</b>
<b>Application</b>	11, 12	13, 26	One – Factor Model: ( $X^2= 4$ , $df= 1$ <b>CFI=0.994</b> , <b>RMSEA= 0.045</b> ) <b>a = 0.51</b>
<b>Questioning</b>	25, 39	24, 37, 40, 42	One – Factor Model: ( $X^2= 30$ , $df= 5$ <b>CFI=0.961</b> , <b>RMSEA= 0.051</b> ) <b>a = 0.39</b>
<b>Classroom as a learning Environment / Teacher – Student Interaction</b>	16, 17	19, 20, 21	One – Factor Model: ( $X^2= 53$ , $df= 5$ <b>CFI=0.943</b> , <b>RMSEA= 0.07</b> ) <b>a = 0.61</b>
<b>Classroom as a learning Environment / Dealing with Misbehaviour</b>	29	27, 33 , 30	One – Factor Model: ( $X^2= 52$ , $df= 2$ <b>CFI=0.951</b> , <b>RMSEA= 0.11</b> ) <b>a = 0.66</b>

**Table 6: Items of the CFA models and results of the SEM and Reliability Analysis for Cyprus**

<b>Teacher Factors</b>	<b>Quantitative Characteristics</b>	<b>Qualitative Characteristics</b>	<b>Results: SEM and Reliability Analysis</b>
<b>Modeling</b>	44, 47	45, 46	One – Factor Model: ( $X^2= 10$ , $df= 2$ <b>CFI=0.999</b> , <b>RMSEA= 0.021</b> ) <b>a = 0.70</b>
<b>Structuring</b>	3, 2, 34, 38	1, 4, 7, 10	Two – Factor Model: ( $X^2= 54$ , $df= 18$ <b>CFI=0.965</b> , <b>RMSEA= 0.033</b> ) <b>a1 =0.47, a2= 0.47</b>
<b>Application</b>	11, 12	13, 26	One – Factor Model: ( $X^2= 2$ , $df= 1$ <b>CFI=0.998</b> , <b>RMSEA= 0.024</b> ) <b>a = 0.56</b>
<b>Questioning</b>	25, 39	24, 37, 40, 42	One – Factor Model: ( $X^2= 62$ , $df= 8$ <b>CFI=0.955</b> , <b>RMSEA= 0.06</b> ) <b>a = 0.51</b>
<b>Classroom as a learning Environment / Teacher – Student Interaction</b>	16, 17	19, 20, 21	One – Factor Model: ( $X^2= 13$ , $df= 5$ <b>CFI=0.993</b> , <b>RMSEA= 0.031</b> ) <b>a = 0.69</b>
<b>Classroom as a learning Environment / Dealing with Misbehaviour</b>	29	27, 33 , 30	One – Factor Model: ( $X^2= 0.23$ , $df= 1$ <b>CFI=0.999</b> , <b>RMSEA= 0.001</b> ) <b>a = 0.71</b>

**Table 7: Items of the CFA models and results of the SEM and Reliability Analysis for Germany**

<b>Teacher Factors</b>	<b>Quantitative Characteristics</b>	<b>Qualitative Characteristics</b>	<b>Results: SEM and Reliability Analysis</b>
<b>Modeling</b>	44, 47	45, 46	One – Factor Model: ( $X^2= 16$ , $df= 2$ <b>CFI=0.998</b> , <b>RMSEA= 0.027</b> ) <b>a = 0.76</b>
<b>Structuring</b>	3, 2, 34, 38	1, 4, 7, 10	Two – Factor Model: ( $X^2= 105$ , $df= 19$ <b>CFI=0.956</b> , <b>RMSEA= 0.06</b> ) <b>a1 =0.72, a2= 0.69</b>
<b>Application</b>	11, 12	13, 26	One – Factor Model: ( $X^2= 1$ , $df= 1$ <b>CFI=0.999</b> , <b>RMSEA= 0.012</b> ) <b>a = 0.65</b>
<b>Questioning</b>	39	24, 37, 40, 42	One – Factor Model: ( $X^2= 169$ , $df= 5$ <b>CFI=0.821</b> , <b>RMSEA= 0.17</b> ) <b>a = 0.45</b>
<b>Classroom as a learning Environment / Teacher – Student Interaction</b>	16, 17	19, 20, 21	One – Factor Model: ( $X^2= 30$ , $df= 5$ <b>CFI=0.972</b> , <b>RMSEA= 0.07</b> ) <b>a = 0.71</b>
<b>Classroom as a learning Environment / Dealing with Misbehaviour</b>	29	27, 33 , 30	One – Factor Model: ( $X^2= 18$ , $df= 2$ <b>CFI=0.981</b> , <b>RMSEA= 0.08</b> ) <b>a = 0.73</b>

**Table 8: Items of the CFA models and results of the SEM and Reliability Analysis for Greece**

<b>Teacher Factors</b>	<b>Quantitative Characteristics</b>	<b>Qualitative Characteristics</b>	<b>Results: SEM and Reliability Analysis</b>
<b>Modeling</b>	44, 47	45, 46	One – Factor Model: ( $X^2= 8$ , $df= 1$ <b>CFI=0.999</b> , <b>RMSEA= 0.028</b> ) <b>a = 0.73</b>
<b>Structuring</b>	3, 2, 34, 38	1, 4, 7, 10	Two – Factor Model: ( $X^2= 51$ , $df= 19$ <b>CFI=0.955</b> , <b>RMSEA= 0.043</b> ) <b>a1 =0.61, a2= 0.45</b>
<b>Application</b>	11, 12	13, 26	One – Factor Model: ( $X^2= 4$ , $df= 1$ <b>CFI=0.991</b> , <b>RMSEA= 0.06</b> ) <b>a = 0.54</b>
<b>Questioning</b>	25, 39	24, 37, 40, 42	One – Factor Model: ( $X^2= 39$ , $df= 8$ <b>CFI=0.945</b> , <b>RMSEA= 0.06</b> ) <b>a = 0.49</b>
<b>Classroom as a learning Environment / Teacher – Student Interaction</b>	16, 17	19, 20, 21	One – Factor Model: ( $X^2= 5$ , $df= 5$ <b>CFI=0.999</b> , <b>RMSEA= 0.011</b> ) <b>a = 0.64</b>
<b>Classroom as a learning Environment / Dealing with Misbehaviour</b>	29	27, 33 , 30	One – Factor Model: ( $X^2= 1$ , $df= 1$ <b>CFI=0.999</b> , <b>RMSEA= 0.27</b> ) <b>a = 0.76</b>

**Table 9: Items of the CFA models and results of the SEM and Reliability Analysis for Ireland**

<b>Teacher Factors</b>	<b>Quantitative Characteristics</b>	<b>Qualitative Characteristics</b>	<b>Results: SEM and Reliability Analysis</b>
<b>Modeling</b>	44, 47	45, 46	One – Factor Model: ( $X^2= 9$ , $df= 2$ <b>CFI=0.999</b> , <b>RMSEA= 0.019</b> ) <b>a = 0.69</b>
<b>Structuring</b>	3, 2, 34, 38	1, 4, 7, 10	Two – Factor Model: ( $X^2= 53$ , $df= 19$ <b>CFI=0.968</b> , <b>RMSEA= 0.029</b> ) <b>a1 =0.51, a2= 0.42</b>
<b>Application</b>	11, 12	13, 26	One – Factor Model: ( $X^2= 0.27$ , $df= 1$ <b>CFI=0.999</b> , <b>RMSEA= 0.001</b> ) <b>a = 0.49</b>
<b>Questioning</b>	25, 39	24, 37, 40, 42	One – Factor Model: ( $X^2= 39$ , $df= 9$ <b>CFI=0.974</b> , <b>RMSEA= 0.040</b> ) <b>a = 0.52</b>
<b>Classroom as a learning Environment / Teacher – Student Interaction</b>	16, 17	19, 20, 21	One – Factor Model: ( $X^2= 45$ , $df= 5$ <b>CFI=0.969</b> , <b>RMSEA= 0.06</b> ) <b>a = 0.65</b>
<b>Classroom as a learning Environment / Dealing with Misbehaviour</b>	29	27, 33 , 30	One – Factor Model: ( $X^2= 0.94$ , $df= 1$ <b>CFI=0.999</b> , <b>RMSEA= 0.001</b> ) <b>a = 0.57</b>

**Table 10: Items of the CFA models and results of the SEM and Reliability Analysis for Slovenia**

<b>Teacher Factors</b>	<b>Quantitative Characteristics</b>	<b>Qualitative Characteristics</b>	<b>Results: SEM and Reliability Analysis</b>
<b>Modeling</b>	44, 47	45, 46	One – Factor Model: ( $X^2= 21$ , $df= 2$ <b>CFI=0.982</b> , <b>RMSEA= 0.070</b> ) <b>a = 0.65</b>
<b>Structuring</b>	3, 2, 34, 38	1, 4, 7, 10	Two – Factor Model: ( $X^2= 139$ , $df= 19$ <b>CFI=0.967</b> , <b>RMSEA= 0.05</b> ) <b>a1 =0.75</b> , <b>a2= 0.69</b>
<b>Application</b>	11, 12	13, 26	One – Factor Model: ( $X^2= 7$ , $df= 1$ <b>CFI=0.995</b> , <b>RMSEA= 0.056</b> ) <b>a = 0.68</b>
<b>Questioning</b>	25, 39	24, 37, 40, 42	One – Factor Model: ( $X^2= 20$ , $df= 8$ <b>CFI=0.990</b> , <b>RMSEA= 0.028</b> ) <b>a = 0.40</b>
<b>Classroom as a learning Environment / Teacher – Student Interaction</b>	16, 17	19, 20, 21	One – Factor Model: ( $X^2= 17$ , $df= 5$ <b>CFI=0.988</b> , <b>RMSEA= 0.035</b> ) <b>a = 0.62</b>
<b>Classroom as a learning Environment / Dealing with Misbehaviour</b>	29	27, 33 , 30	One – Factor Model: ( $X^2= 62$ , $df= 2$ <b>CFI=0.97</b> , <b>RMSEA= 0.12</b> ) <b>a = 0.73</b>

## 2.3 Within Countries SEM analysis

For the within countries analysis our assumption was again in line with the main assumptions of the dynamic model. More specifically, our assumption was that the factors of management of time, teacher ability to deal with student misbehaviour and questioning: raising non-appropriate questions belong to a second order factor, whereas the other factors belong to another second order factor. Therefore we have conducted separate SEM analysis for each country to see whether the grouping of factors was similar to our assumptions. From the within country SEM analysis, it was found out that the model that was produced by the across country analysis fits well to the data emerged by each country separately. Specifically, two second order factors have again been identified. The first second order factor is consisted of the factors measuring management of time, teacher ability to deal with student misbehaviour and questioning: raising non-appropriate questions. All the other factors were found to load on the second overarching factor which can indicate the quality of teaching. It is important to note that the separate within country SEM analysis showed that the correlation in each country between these two overarching factors is very small which implies again that those teachers who are able to use the teaching time are not necessarily able to use effectively the teaching time.

Also, the two other models that were tested in the across countries analysis were tested in the within country analysis to compare their fit to the data with the proposed model. In the first model (Model 2) again all the items that were used for the SEM analysis were considered as belonging to a single factor, to test again whether the questionnaire items refer to a social desirability factor and thereby the questionnaire is not valid. In the second model (Model 3) the items concerning quality of teaching were considered as one factor and the items concerning quantity of teaching as another factor. The fit indices of each model are also shown in Tables 11-16. We can see that model 1 best fits the data in each country and that its fit indices were very good.

The results of the SEM analysis per country are presented in more detail below:

**1) Belgium:**

**Table 11: Results of the SEM analysis in Belgium**

<b>SEM analysis – Results</b>							
<b>Models for Belgium (N= 1908)</b>	<b>X<sup>2</sup></b>	<b>Df</b>	<b>X<sup>2</sup>/df</b>	<b>P</b>	<b>CFI</b>	<b>RMSEA</b>	<b>Range RMSEA</b>
<b>Model 1</b>	731	297	2,4	0.01	0.929	0.028	0.025 - 0.030
<b>Model 2</b>	2668	324	8,3	0.001	0.616	0.061	0.059 – 0.064
<b>Model 3</b>	1395	323	4,3	0.001	0.824	0.042	0.039 – 0.044

As can be seen in Table 11, model 1 which is the same as the model which was presented in Figure 1 and was produced from the across country analysis, is the model that best fits the data in Belgium and its fit indices are very high. In addition, the loadings of the items and the factors are also high and all the loadings are statistically significant. The only exception is item 25 (V28) where the loading is relatively low. The loadings are presented in Figure 1 in Appendix D. Also, in order for the model to best fit the data of the country, item 5 from assessment factor was removed from the model as its loading was low and therefore question 9 (V34) is regressed to the second order factor. From Figure 1 (Appendix D) we can also see that the correlation between the two overarching factors in Belgium is very small which suggests that effectiveness in one overarching factor does not also imply effectiveness in the other.

Based on the loadings of the items that occurred from the SEM analysis, as they appear in Figure 1 (Appendix D) the factor scores which will be used for the multilevel analysis were produced.



## 2) Cyprus:

**Table 12: Results of the SEM analysis in Cyprus**

SEM analysis – Results							
Models for Cyprus (N= 1881)	X <sup>2</sup>	Df	X <sup>2</sup> /df	P	CFI	RMSEA	Range RMSEA
<b>Model 1</b>	825	317	2,6	0.01	0.943	0.029	0.027 - 0.032
<b>Model 2</b>	3441	350	9,8	0.001	0.652	0.069	0.066 – 0.071
<b>Model 3</b>	1584	349	4,3	0.001	0.861	0.043	0.041 – 0.046

In Table 12, the fit indices that were produced from the SEM analysis in Cyprus are presented and as we can see, model 1 is the model that best fits the data and its fit indices are very high. Model 1 is the same as the model which was presented in Figure 1 and was produced from the across country analysis. In addition, the loadings of the items and the factors are also high and all the loadings are statistically significant. The structure of the model for Cyprus is the same as the across countries model and no further items had to be removed. The loadings are presented in Figure 2 in Appendix D. From Figure 2 (Appendix D) we can also see that the correlation between the two overarching factors of Quality and Quantity of teaching in Cyprus is very small.

Based on the loadings of the items that occurred from the SEM analysis, as they appear in Figure 2 (Appendix D) the factor scores which will be used for the multilevel analysis were produced.

### 3) Greece

**Table 13: Results of the SEM analysis in Greece**

SEM analysis – Results							
Models for Greece (N= 905)	X <sup>2</sup>	Df	X <sup>2</sup> /df	P	CFI	RMSEA	Range RMSEA
<b>Model 1</b>	560	312	1,8	0.01	0.944	0.030	0.026 - 0.034
<b>Model 2</b>	2386	350	6,8	0.001	0.542	0.080	0.077 – 0.083
<b>Model 3</b>	1285	349	3,7	0.001	0.789	0.054	0.051 – 0.058

As can be seen in Table 13, model 1 which is the same as the model which was presented in Figure 1 and was produced from the across country analysis, is the model that best fits the data in Greece and its fit indices are very high. In addition, the loadings of the items and the factors are also high and all the loadings are statistically significant. Like Belgium, the only exception is item 25 (V28) where the loading is relatively low. The structure of the model for Greece is the same as the across countries model and no further items had to be removed. The loadings are presented in Figure 3 in Appendix D. The SEM results for Greece also showed a very small correlation between the two overarching factors of Quality and Quantity of teaching, which is in line with the results of the other countries and the across country analysis.

Based on the loadings of the items that occurred from the SEM analysis, as they appear in Figure 3 (Appendix D) the factor scores which will be used for the multilevel analysis were produced.

#### 4) Ireland

**Table 14: Results of the SEM analysis in Ireland**

SEM analysis – Results							
Models for Ireland (N= 2140)	X <sup>2</sup>	Df	X <sup>2</sup> /df	P	CFI	RMSEA	Range RMSEA
<b>Model 1</b>	915	327	2,8	0.01	0.929	0.029	0.027 - 0.031
<b>Model 2</b>	2416	350	6,9	0.001	0.752	0.053	0.051 – 0.055
<b>Model 3</b>	1416	349	4,1	0.001	0.872	0.038	0.036 – 0.040

The SEM analysis conducted for Ireland showed again that model 1 which is the same as the model which was presented in Figure 1 and was produced from the across country analysis, is the model that best fits the data in Ireland and its fit indices are very high. In addition, the loadings of the items and the factors are also high and all the loadings are statistically significant. The structure of the model for Ireland is the same as the across countries model and no further items had to be removed. The loadings are presented in Figure 4 in Appendix D. The SEM results for Ireland showed that the correlation between the two overarching factors of Quality and Quantity of teaching is relatively high (0.43), which shows that teachers' effectiveness in one overarching factor depends on their effectiveness in the other.

Based on the loadings of the items that occurred from the SEM analysis, as they appear in Figure 4 (Appendix D) the factor scores which will be used for the multilevel analysis were produced.

## 5) Slovenia

**Table 15: Results of the SEM analysis in Slovenia**

SEM analysis – Results							
Models for Slovenia (N= 2049)	X <sup>2</sup>	Df	X <sup>2</sup> /df	P	CFI	RMSEA	Range RMSEA
<b>Model 1</b>	1158	281	4,1	0.01	0.926	0.039	0.037 - 0.041
<b>Model 2</b>	4573	324	14,1	0.001	0.640	0.080	0.078 – 0.082
<b>Model 3</b>	2196	323	6,8	0.001	0.841	0.053	0.051 – 0.055

As it is shown in Table 15, model 1 is the model that best fits the data in Slovenia in comparison to the other two models and its fit indices are very high. Model 1 is the same as the model which was presented in Figure 1 and was produced from the across country analysis. In addition, the loadings of the items and the factors are also high and all the loadings are statistically significant. The loadings are presented in Figure 5 in Appendix D. In order for the model to best fit the data of the country, item 10 from the factor of structuring (qualitative characteristics) was excluded from the analysis of Slovenia as its loading was low. From Figure 5 (Appendix D) we can also see that the correlation between the two overarching factors in Slovenia is very small which is in line with the results of the across country analysis.

Based on the loadings of the items that occurred from the SEM analysis, as they appear in Figure 5 (Appendix D) the factor scores which will be used for the multilevel analysis were produced.

## 6) Germany

**Table 16: Results of the SEM analysis in Germany**

SEM analysis – Results							
Models for Germany (N= 1072)	X <sup>2</sup>	Df	X <sup>2</sup> /df	P	CFI	RMSEA	Range RMSEA
<b>Model 1</b>	547	219	2,5	0.01	0.959	0.037	0.034 - 0.041
<b>Model 2</b>	3472	275	12,6	0.001	0.599	0.104	0.101 – 0.107
<b>Model 3</b>	1434	274	5,2	0.001	0.855	0.063	0.060 – 0.066

As noted earlier, item 25 was not administered in Germany item and therefore could not be used for the country analyses.

From Table 16 we can see that model 1, which is the same as the model which was presented in Figure 1 and was produced from the across country analysis, is the model that best fits the data in Germany and its fit indices are very high. In addition, the results of the SEM analysis showed that the loadings of the items and the factors are also high and all the loadings are statistically significant. The only exception is item 1 (V9) from the qualitative characteristics of structuring where the loading is relatively low. The loadings are presented in Figure 6 in Appendix D. In order for the model to best fit the data of the country, items 9 from assessment and 17 from classroom as a learning environment: teacher-student interaction were excluded from the analysis in Germany as their loadings were found to be low. Therefore question 5 (V34) is regressed on the second order factor. From Figure 6 (Appendix D) we can also see that the correlation between the two overarching factors in Germany is also is very small.

Based on the loadings of the items that occurred from the SEM analysis, as they appear in Figure 6 (Appendix D) the factor scores which will be used for the multilevel analysis were produced.

## **References**

- Barnette, J. J. (2000). Effects of stem and Likert response option reversals on survey internal consistency: If you feel the need, there is a better alternative to using those negatively worded stems. *Educational and Psychological Measurement, 60*, 361–370.
- Bryne, B. M., (1994). *Structural Equation Modeling with EQS and EQS/Windows*. Newbury Park, CA: Sage.
- Cronbach, L.J. (3rd Ed) (1990). *Essentials of Psychological Testing*. New York: Harper & Row.
- Creemers, B.P.M., & Kyriakides, L. (2008). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools*. London: Routledge.
- Creemers, B.P.M., & Kyriakides, L. (2010). Explaining stability and changes in school effectiveness by looking at changes in the functioning of school factors. *School Effectiveness and School Improvement, 21*(4), 409–427.
- Kyriakides, L., & Creemers, B.P.M. (2008). Using a multidimensional approach to measure the impact of classroom-level factors upon student achievement: a study testing the validity of the dynamic model. *School Effectiveness and School Improvement, 19*(2), 183–205.

## **Appendix A**

**A1) Table 1: Missing values across countries**

Items of the Student Questionnaire	Total Across Countries		
	Code 7	Code 8	Code 9
Q1	139	13	82
Q2	139	10	85
Q3	139	15	120
Q4	139	14	141
Q5	139	4	103
Q6	139	24	180
Q7	139	10	146
Q8	139	15	140
Q9	139	8	125
Q10	139	13	161
Q11	139	9	168
Q12	139	10	173
Q13	139	11	193
Q14	139	14	167
Q15	139	15	144
Q16	139	8	136
Q17	139	10	128
Q18	139	14	185
Q19	139	10	235
Q20	139	17	213
Q21	139	11	187
Q22	139	17	222
Q23	139	14	215
Q24	139	16	215
Q25	139	15	1301
Q26	139	26	217
Q27	139	12	208
Q28	139	8	230
Q29	139	14	232
Q30	139	24	262
Q31	139	24	280
Q32	139	16	266
Q33	139	14	286
Q34	139	11	290
Q35	139	18	318
Q36	139	13	308

Q37	139	13	438
Q38	139	14	292
Q39	139	17	293
Q40	139	13	337
Q41	139	15	322
Q42	139	15	329
Q43	139	11	385
Q44	139	14	335
Q45	139	10	334
Q46	139	9	367
Q47	139	14	387
Q48	139	10	400
Q49	139	7	349

**A2) Table 2: Missing values in Belgium**

Items of the Student Questionnaire	Belgium		
	Code 7	Code 8	Code 9
Q1	139	0	7
Q2	139	2	5
Q3	139	0	4
Q4	139	2	6
Q5	139	1	5
Q6	139	2	24
Q7	139	1	20
Q8	139	3	6
Q9	139	2	9
Q10	139	5	10
Q11	139	1	15
Q12	139	1	6
Q13	139	3	0
Q14	139	2	5
Q15	139	3	9
Q16	139	1	10
Q17	139	1	6
Q18	139	1	13
Q19	139	0	12
Q20	139	1	14
Q21	139	2	10
Q22	139	1	5
Q23	139	0	12



Q24	139	5	8
Q25	139	2	6
Q26	139	2	8
Q27	139	1	12
Q28	139	5	11
Q29	139	3	10
Q30	139	2	9
Q31	139	0	8
Q32	139	1	6
Q33	139	1	5
Q34	139	0	8
Q35	139	2	14
Q36	139	0	10
Q37	139	1	21
Q38	139	2	12
Q39	139	3	17
Q40	139	3	22
Q41	139	2	13
Q42	139	2	11
Q43	139	2	13
Q44	139	3	5
Q45	139	2	7
Q46	139	1	16
Q47	139	2	19
Q48	139	1	14
Q49	139	1	7

**A3) Table 3: Missing values in Cyprus**

Items of the Student Questionnaire	Cyprus		
	Code 7	Code 8	Code 9
Q1	0	0	19
Q2	0	0	23
Q3	0	0	34
Q4	0	0	31
Q5	0	0	18
Q6	0	0	20
Q7	0	0	28
Q8	0	0	33
Q9	0	0	34
Q10	0	0	48

Q11	0	0	47
Q12	0	0	52
Q13	0	0	66
Q14	0	0	45
Q15	0	0	44
Q16	0	0	30
Q17	0	0	28
Q18	0	0	47
Q19	0	0	57
Q20	0	0	52
Q21	0	0	56
Q22	0	0	64
Q23	0	0	57
Q24	0	0	69
Q25	0	0	82
Q26	0	0	61
Q27	0	0	52
Q28	0	0	65
Q29	0	0	72
Q30	0	0	91
Q31	0	0	87
Q32	0	0	84
Q33	0	0	95
Q34	0	0	83
Q35	0	0	98
Q36	0	0	100
Q37	0	0	83
Q38	0	0	80
Q39	0	0	73
Q40	0	0	85
Q41	0	0	99
Q42	0	0	100
Q43	0	0	113
Q44	0	0	100
Q45	0	0	103
Q46	0	0	109
Q47	0	0	112
Q48	0	0	111
Q49	0	0	104

**A4) Table 4: Missing values in Germany**

Items of the Student Questionnaire	Germany		
	Code 7	Code 8	Code 9
Q1	0	2	19
Q2	0	1	17
Q3	0	1	24
Q4	0	1	34
Q5	0	1	26
Q6	0	1	42
Q7	0	1	20
Q8	0	2	34
Q9	0	2	30
Q10	0	1	30
Q11	0	3	28
Q12	0	1	30
Q13	0	1	34
Q14	0	2	29
Q15	0	1	22
Q16	0	1	26
Q17	0	4	28
Q18	0	8	40
Q19	0	3	44
Q20	0	2	28
Q21	0	2	28
Q22	0	3	44
Q23	0	4	49
Q24	0	1	36
Q25	0	0	1072
Q26	0	3	40
Q27	0	2	41
Q28	0	3	43
Q29	0	2	34
Q30	0	2	33
Q31	0	9	44
Q32	0	4	45
Q33	0	6	45
Q34	0	4	47
Q35	0	3	49
Q36	0	0	38
Q37	0	2	42
Q38	0	4	48

Q39	0	2	48
Q40	0	3	51
Q41	0	4	46
Q42	0	2	50
Q43	0	2	46
Q44	0	3	55
Q45	0	4	52
Q46	0	3	54
Q47	0	4	59
Q48	0	4	62
Q49	0	2	55

**A5) Table 5: Missing values in Greece**

	Greece		
Items of the Student Questionnaire	Code 7	Code 8	Code 9
Q1	0	0	12
Q2	0	0	9
Q3	0	0	17
Q4	0	0	21
Q5	0	0	13
Q6	0	0	18
Q7	0	0	29
Q8	0	0	31
Q9	0	0	24
Q10	0	0	23
Q11	0	0	32
Q12	0	0	41
Q13	0	0	47
Q14	0	0	30
Q15	0	0	37
Q16	0	0	45
Q17	0	0	38
Q18	0	0	45
Q19	0	0	67
Q20	0	0	55
Q21	0	0	57
Q22	0	0	74
Q23	0	0	66
Q24	0	0	72
Q25	0	0	82

Q26	0	0	68
Q27	0	0	65
Q28	0	0	79
Q29	0	0	86
Q30	0	0	82
Q31	0	0	96
Q32	0	0	93
Q33	0	0	100
Q34	0	0	108
Q35	0	0	116
Q36	0	0	110
Q37	0	0	108
Q38	0	0	110
Q39	0	0	111
Q40	0	0	116
Q41	0	0	137
Q42	0	0	125
Q43	0	0	143
Q44	0	0	129
Q45	0	0	131
Q46	0	0	126
Q47	0	0	134
Q48	0	0	131
Q49	0	0	132

**A6) Table 6: Missing values in Ireland**

	Ireland		
Items of the Student Questionnaire	Code 7	Code 8	Code 9
Q1	0	0	20
Q2	0	0	19
Q3	0	0	33
Q4	0	0	34
Q5	0	0	33
Q6	0	0	48
Q7	0	0	23
Q8	0	0	14
Q9	0	0	18
Q10	0	0	27
Q11	0	0	32
Q12	0	0	29

Q13	0	0	38
Q14	0	0	50
Q15	0	0	18
Q16	0	0	17
Q17	0	0	21
Q18	0	0	31
Q19	0	0	40
Q20	0	0	54
Q21	0	0	20
Q22	0	0	19
Q23	0	0	16
Q24	0	0	16
Q25	0	0	29
Q26	0	0	24
Q27	0	0	24
Q28	0	0	21
Q29	0	0	11
Q30	0	0	15
Q31	0	0	27
Q32	0	0	25
Q33	0	0	28
Q34	0	0	33
Q35	0	0	28
Q36	0	0	31
Q37	0	0	16
Q38	0	0	24
Q39	0	0	23
Q40	0	0	44
Q41	0	0	14
Q42	0	0	22
Q43	0	0	41
Q44	0	0	38
Q45	0	0	31
Q46	0	0	35
Q47	0	0	42
Q48	0	0	68
Q49	0	0	40

**A7) Table 7: Missing values in Slovenia**

Items of the Student Questionnaire	Slovenia		
	Code 7	Code 8	Code 9
Q1	0	11	5
Q2	0	7	12
Q3	0	14	8
Q4	0	11	16
Q5	0	2	8
Q6	0	21	28
Q7	0	8	26
Q8	0	10	22
Q9	0	4	10
Q10	0	7	23
Q11	0	5	14
Q12	0	8	15
Q13	0	7	8
Q14	0	10	8
Q15	0	11	14
Q16	0	6	8
Q17	0	5	7
Q18	0	5	9
Q19	0	7	15
Q20	0	14	10
Q21	0	6	16
Q22	0	13	16
Q23	0	10	15
Q24	0	10	14
Q25	0	13	30
Q26	0	21	16
Q27	0	9	15
Q28	0	0	11
Q29	0	9	19
Q30	0	20	32
Q31	0	15	18
Q32	0	11	13
Q33	0	7	13
Q34	0	7	11
Q35	0	13	13
Q36	0	13	19
Q37	0	10	16
Q38	0	8	18

Q39	0	12	21
Q40	0	7	19
Q41	0	9	13
Q42	0	11	21
Q43	0	7	29
Q44	0	8	8
Q45	0	4	10
Q46	0	5	27
Q47	0	8	21
Q48	0	5	14
Q49	0	4	11

**A8) Table 8: Results of the ANOVA analysis across countries**

	F	Sig.
V1	2,980	,000
V2	2,469	,000
V3	2,791	,000
V4	3,596	,000
V5	3,264	,000
V6	3,386	,000
V7	2,515	,000
V8	2,296	,000
V9	2,620	,000
V10	2,513	,000
V11	2,693	,000
V12	2,457	,000
V13	2,597	,000
V14	4,977	,000
V15	4,621	,000
V16	1,731	,000
V17	3,148	,000
V18	3,398	,000
V19	2,355	,000
V20	2,308	,000
V21	3,854	,000
V22	3,365	,000
V23	6,143	,000
V24	1,709	,000
V25	2,596	,000
V26	2,352	,000
V27	4,505	,000
V28	3,048	,000



V29	3,346	,000
V30	4,348	,000
V31	4,183	,000
V32	4,420	,000
V33	3,711	,000
V34	2,899	,000
V35	2,913	,000
V36	3,274	,000
V37	3,040	,000
V38	2,480	,000
V39	2,082	,000
V40	2,230	,000
V41	2,981	,000
V42	2,101	,000
V43	3,027	,000
V44	2,392	,000
V45	2,971	,000
V46	2,592	,000
V47	3,938	,000
V48	3,305	,000
V49	1,888	,000

## Appendix B

# SURVEY OF PRIMARY SCHOOLS

Dear student,

The [NAME OF INSTITUTION] carries out a lot of research including research on education.

We are conducting a study on students in Grade 4 and would like to know your opinion about the teaching of Mathematics and Science in your classroom.

**The answers you give will not be shown to your teachers, anyone else in your school or your parents.**

**We are giving each student a special number so you do not need to write your name on the questionnaire.**

Please answer **all** of the questions. To answer the questions, please circle a number on each line.

Please ask the interviewer if you do not understand what to do.

## PART A

After each statement you read there are four numbers. Think carefully and put a circle around the number that most fits your opinion:

1: if this **never** happens in your class

2: if this **rarely** happens in your class

3: if this **sometimes** happens in your class

4: if this **often** happens in your class

		Never	Rarely	Sometimes	Often	Almost Always
Q1.	In Mathematics and Science, we start the lesson with things that are easy to understand. As the lesson goes on what we cover is more difficult.	1	2	3	4	5
Q2.	The teacher gives us exercises at the beginning of the lesson to check what we have learnt from the previous lesson.	1	2	3	4	5
Q3.	At the beginning of the lesson, the teacher starts with what we covered in the previous lessons.	1	2	3	4	5
Q4.	My teacher helps us to understand how different activities (such as exercises, subject matter) during a lesson are related to each other.	1	2	3	4	5
Q5.	A few days before the test, my teacher gives us similar exercises to those that will be in the test.	1	2	3	4	5
Q6.	My teacher tells my parents how good I am compared to my classmates when they visit her/him (or in my school report).	1	2	3	4	5
Q7.	When the teacher is teaching, I always know what part of the lesson (beginning, middle, end) we are in.	1	2	3	4	5

After each statement you read there are four numbers. Think carefully and put a circle around the number that most fits your opinion:

- 1: if the situation described **never** happens in your class
- 2: if the situation described happens **rarely** in your class
- 3: if the situation described happens **sometimes** in your class
- 4: if the situation described happens **often** in your class

		Never	Rarely	Sometimes	Often	Almost Always
Q8.	When doing an activity in Mathematics and Science I know why I am doing it.	1	2	3	4	5
Q9.	When we go over our homework, our teacher finds what we had problems with and helps us to overcome these difficulties.	1	2	3	4	5
Q10.	Our teacher has good ways of explaining how the new things we are learning are related to things we already know.	1	2	3	4	5
Q11.	At the end of each lesson, the teacher gives us exercises on what we have just learned.	1	2	3	4	5
Q12.	During the lesson our teacher often covers the same things that we have already learned or done exercises in.	1	2	3	4	5
Q13.	The teacher immediately comes to help me when I have problems doing an activity.	1	2	3	4	5
Q14.	The teacher gives more exercises to some pupils than the rest of the class.	1	2	3	4	5
Q15.	The teacher gives some pupils different exercises to do than the rest of the class.	1	2	3	4	5

After each statement you read there are four numbers. Think carefully and put a circle around the number that most fits your opinion:

- 1: if the situation described **never** happens in your class
- 2: if the situation described happens **rarely** in your class
- 3: if the situation described happens **sometimes** in your class
- 4: if the situation described happens **often** in your class

		Never	Rarely	Sometimes	Often	Almost Always
Q16.	The teacher gives all pupils the chance to take part in the lesson.	1	2	3	4	5
Q17.	Our teacher encourages us to work together with our classmates during Mathematics and Science lessons.	1	2	3	4	5
Q18.	Some pupils in my classroom work together when our teacher asks us but some pupils do not.	1	2	3	4	5
Q19.	Our teacher makes us feel that we can ask him/her for help or advice if we need it.	1	2	3	4	5
Q20.	Our teacher encourages us to ask questions if there is something that we do not understand during the lesson.	1	2	3	4	5
Q21.	During the lesson, our teacher encourages and tells us that we are doing good work (i.e. she/he says to us "well done").	1	2	3	4	5
Q22.	When we are working in teams, our teacher encourages competition between teams. (If you do not work in teams, please circle the number 1).	1	2	3	4	5

After each statement you read there are four numbers. Think carefully and put a circle around the number that most fits your opinion:

- 1: if the situation described **never** happens in your class
- 2: if the situation described happens **rarely** in your class
- 3: if the situation described happens **sometimes** in your class
- 4: if the situation described happens **often** in your class

		Never	Rarely	Sometimes	Often	Almost Always
Q23.	In Mathematics and Science lessons, some of my classmates hide their work and answers so that none of the other pupils can see it.	1	2	3	4	5
Q24.	When a pupil gives a wrong answer the teacher helps her/him to understand her/his mistake and find the correct answer.	1	2	3	4	5
Q25.	When the teacher asks us a question about the lesson he/she asks us for the answer but does not ask us to explain how we worked out the answer.	1	2	3	4	5
Q26.	When one of the pupils in the class is having difficulties with the lesson, our teacher goes to help him/her straight away.	1	2	3	4	5
Q27.	There are some pupils in the classroom that tease some of their classmates during Mathematics and Science lessons.	1	2	3	4	5
Q28.	I know that if I break a class rule I will be punished.	1	2	3	4	5
Q29.	The teacher has to stop teaching the class because one of the pupils is being naughty	1	2	3	4	5

After each statement you read there are four numbers. Think carefully and put a circle around the number that most fits your opinion:

- 1: if the situation described **never** happens in your class
- 2: if the situation described happens **rarely** in your class
- 3: if the situation described happens **sometimes** in your class
- 4: if the situation described happens **often** in your class

		Never	Rarely	Sometimes	Often	Almost Always
Q30.	When a pupil gives a wrong answer in Mathematics and Science class some of the other children in the class make fun of her/him.	1	2	3	4	5
Q31.	Our teacher keeps on teaching us even though it is break-time or the lesson is supposed to be over.	1	2	3	4	5
Q32.	When I finish a task before my classmates my teacher immediately gives me something else to do.	1	2	3	4	5
Q33.	When the teacher talks to a pupil after they have been naughty, sometimes after a while, that pupil will be naughty again.	1	2	3	4	5
Q34.	We spend time at the end of the lesson to go over what we have just learned.	1	2	3	4	5
Q35.	There are times we do not have the necessary materials for the lesson to take place (e.g., dienes, unifix, test tubes, thermometers, calculators, rulers)	1	2	3	4	5
Q36.	There are times when I do not have anything to do during a lesson.	1	2	3	4	5

After each statement you read there are four numbers. Think carefully and put a circle around the number that most fits your opinion:

- 1: if the situation described **never** happens in your class
- 2: if the situation described happens **rarely** in your class
- 3: if the situation described happens **sometimes** in your class
- 4: if the situation described happens **often** in your class

		Never	Rarely	Sometimes	Often	Almost Always
Q37.	During a Mathematics or Science lesson, our teacher asks us to give our own opinion on a certain issue.	1	2	3	4	5
Q38.	Our teacher asks us questions at the beginning of the lesson to help us remember what we did in the previous lesson.	1	2	3	4	5
Q39.	Our teacher uses words that are hard to understand when he/she asks us a question.	1	2	3	4	5
Q40.	When we do not understand a question, our teacher says it in a different way so we can understand it.	1	2	3	4	5
Q41.	When a pupil gives a wrong answer our teacher gets another pupil to answer the question.	1	2	3	4	5
Q42.	When I give a wrong answer to a question the teacher helps me to understand my mistake and find the correct answer.	1	2	3	4	5
Q43.	Our teacher praises all pupils the same when we answer a question correctly.	1	2	3	4	5



After each statement you read there are four numbers. Think carefully and put a circle around the number that most fits your opinion:

- 1: if the situation described **never** happens in your class
- 2: if the situation described happens **rarely** in your class
- 3: if the situation described happens **sometimes** in your class
- 4: if the situation described happens **often** in your class

		Never	Rarely	Sometimes	Often	Almost Always
Q44.	When we have problem solving exercises and tasks in Mathematics and Science lessons, our teacher helps us by showing us easy ways or tricks to solve the exercises or tasks.	1	2	3	4	5
Q45.	Our teacher lets us use our own easy ways or tricks to solve the exercises or tasks we have in Mathematics and Science.	1	2	3	4	5
Q46.	In Mathematics and Science lessons, our teacher teaches us ways or tricks that can be used in different lessons.	1	2	3	4	5
Q47.	Our teacher encourages us to find ways or tricks to solve the exercises or work s/he gives us.	1	2	3	4	5
Q48.	I am there when my teacher talks to my parents for my progress.	1	2	3	4	5
Q49.	When we are having a test I finish up within the time given to us.	1	2	3	4	5

**PART B**

In this part there are some statements. For each statement circle the answer that shows what usually happens in your class during Mathematics and Science lessons.

We have tests

- A. Every week
- B. Every two weeks
- C. Every month
- D. Every term
- E. Never

The teacher gives corrected tests back to us

- A. Within a week
- B. Within two weeks
- C. Within three weeks
- D. In a month or even longer
- E. S/he never returns them.

The teacher explains to us what s/he expects us to learn from the Mathematics and Science lessons. This happens:

- A. in every lesson
- B. in most of the lessons
- C. only sometimes
- D. very rarely
- E. never.

When no student raises his/her hand to answer a question, the teacher usually (please choose one answer)

- A. answers the question and moves to something else
- B. repeats the question using the same words
- C. restates the question using simpler words
- D. asks an easier question
- E. gives us hints or clues to help us answer the question.

Further below, write down any comments you want to make about the questionnaire and about the teaching of mathematics and science in your classroom.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

**Thank you for your cooperation**

## **Appendix C**

**Specification Table: Items of the students' questionnaire by factor and dimension**

	<b>Dimensions</b>				
<b>Teacher Factors</b>	<b>Frequency</b>	<b>Focus</b>	<b>Stage</b>	<b>Quality</b>	<b>Differentiation</b>
<b>Orientation</b>				8	
<b>Structuring</b>	3	10	2, 34, 38	1, 4, 7	
<b>Application</b>			11, 12	26	13, 14, 15, 32
<b>Management of Time</b>	31, 35, 36	Not applicable (N/A)			
<b>Questioning</b>	25, 39			24, 37, 40, 41, 42	43
<b>Modeling</b>	44, 47			45, 46	
<b>Classroom as a learning Environment / Teacher – Student Interaction</b>	16, 17			19, 20, 21, 22	
<b>Classroom as a learning Environment / Dealing with Misbehaviour</b>	29, 18	28		23, 27, 33, 30	
<b>Assessment</b>	50, 51			5, 6, 9, 48, 49	

## **Appendix D**

### **Figures with the second-order factor model of the student questionnaire measuring teacher factors with factor parameter estimates for the within country analysis**

#### **First Order Factors:**

**F1:** Modeling

**F2:** Structuring – Quantitative Characteristics

**F3:** Structuring – Qualitative Characteristics

**F4:** Application

**F6:** Management of Time

**F7:** Classroom as a Learning Environment – Qualitative characteristics: Teacher - Student interaction

**F8:** Classroom as a Learning Environment – Quantitative Characteristics: Dealing with Misbehaviour

**F9:** Questioning – Quantitative Characteristics: Raising non-appropriate questions

**F10:** Questioning – Qualitative Characteristics

**F11:** Assessment

**V1:** Orientation

#### **Second Order Factors:**

**SF1:** Quality of Teaching

**SF2:** Quantity of Teaching (Management of Time, Misbehaviour & Questioning: raising non – appropriate questions)

Figure 1: SEM results for Belgium

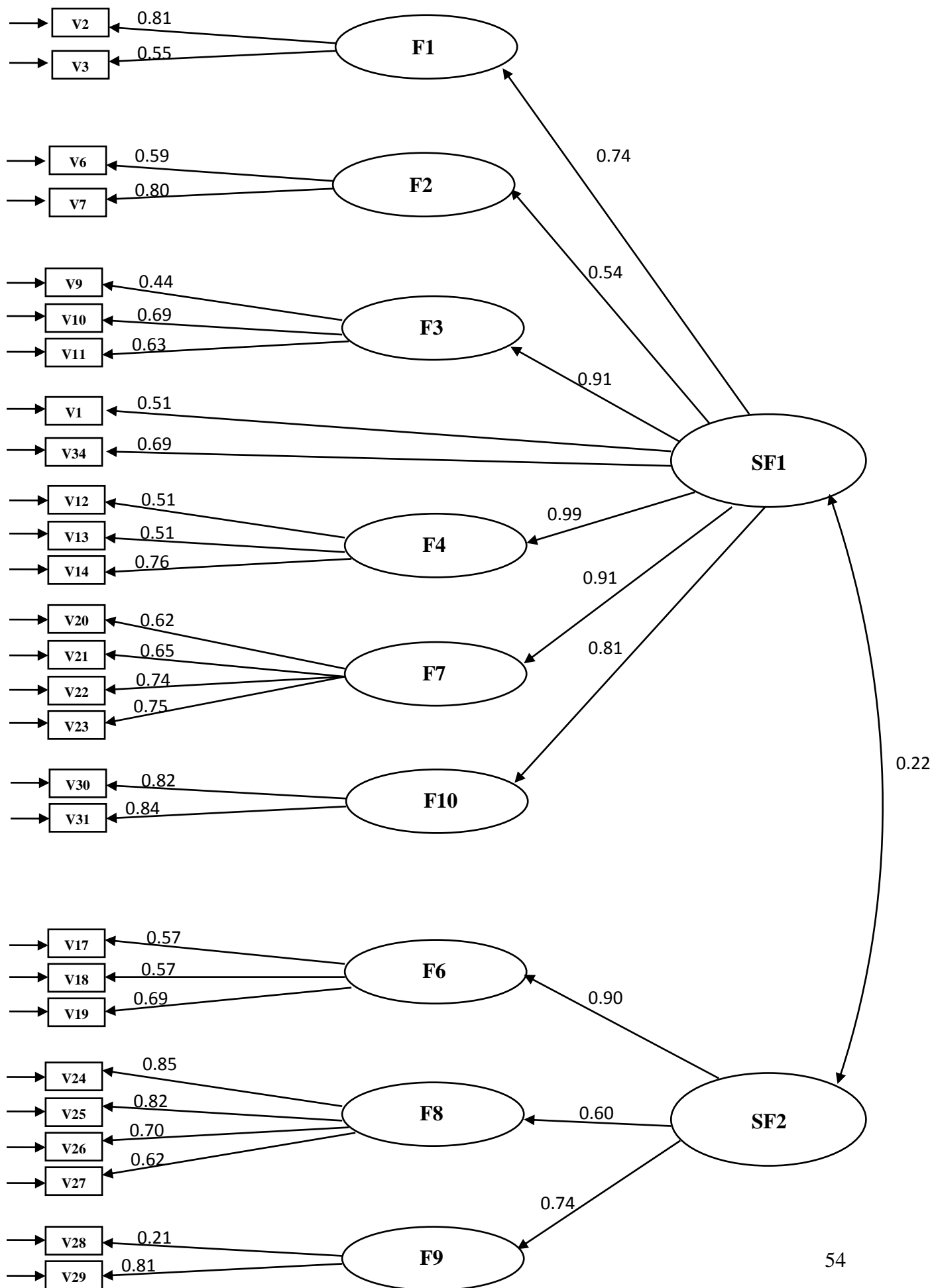


Figure2: SEM results for Cyprus

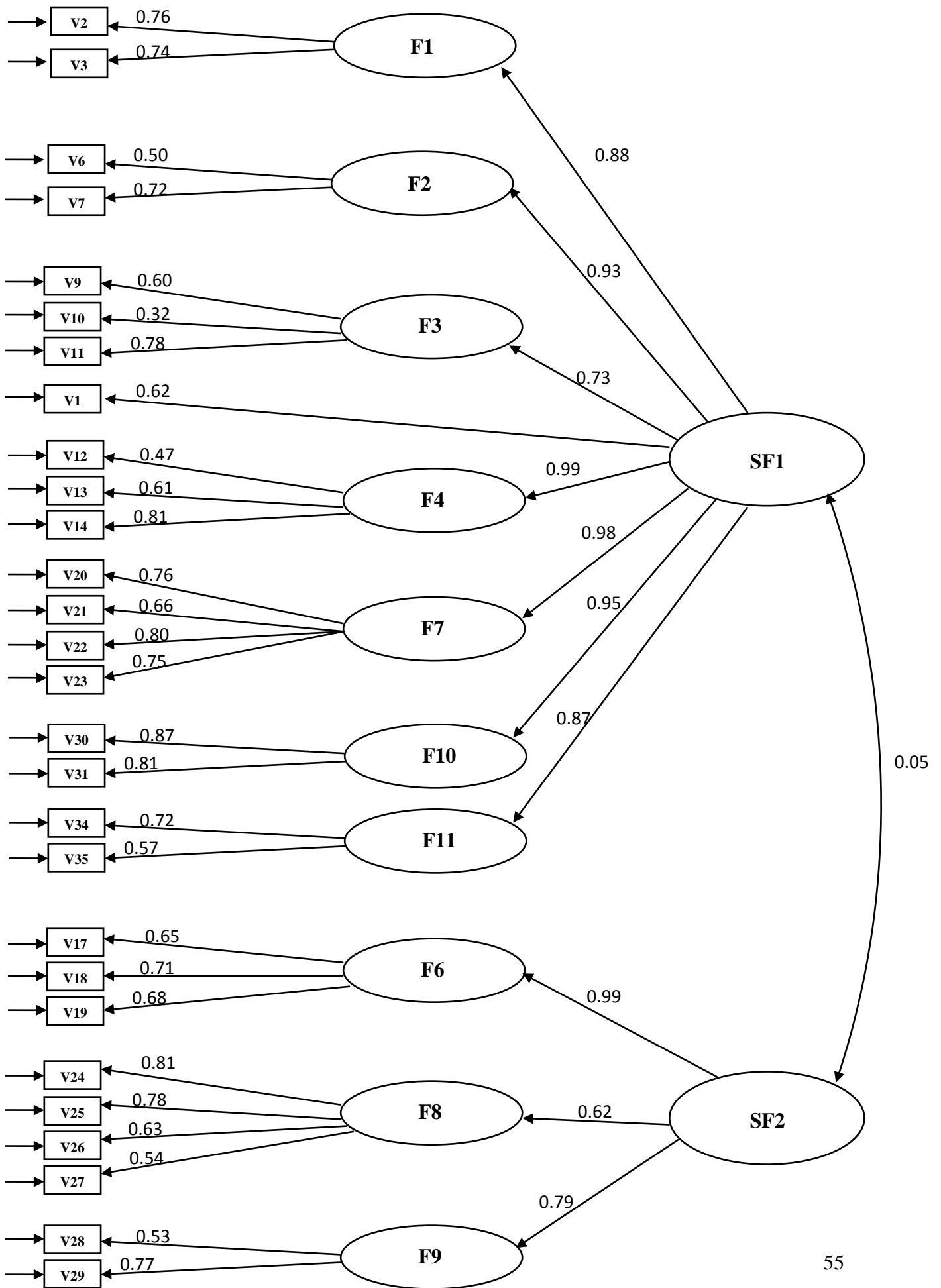


Figure 3: SEM results for Greece

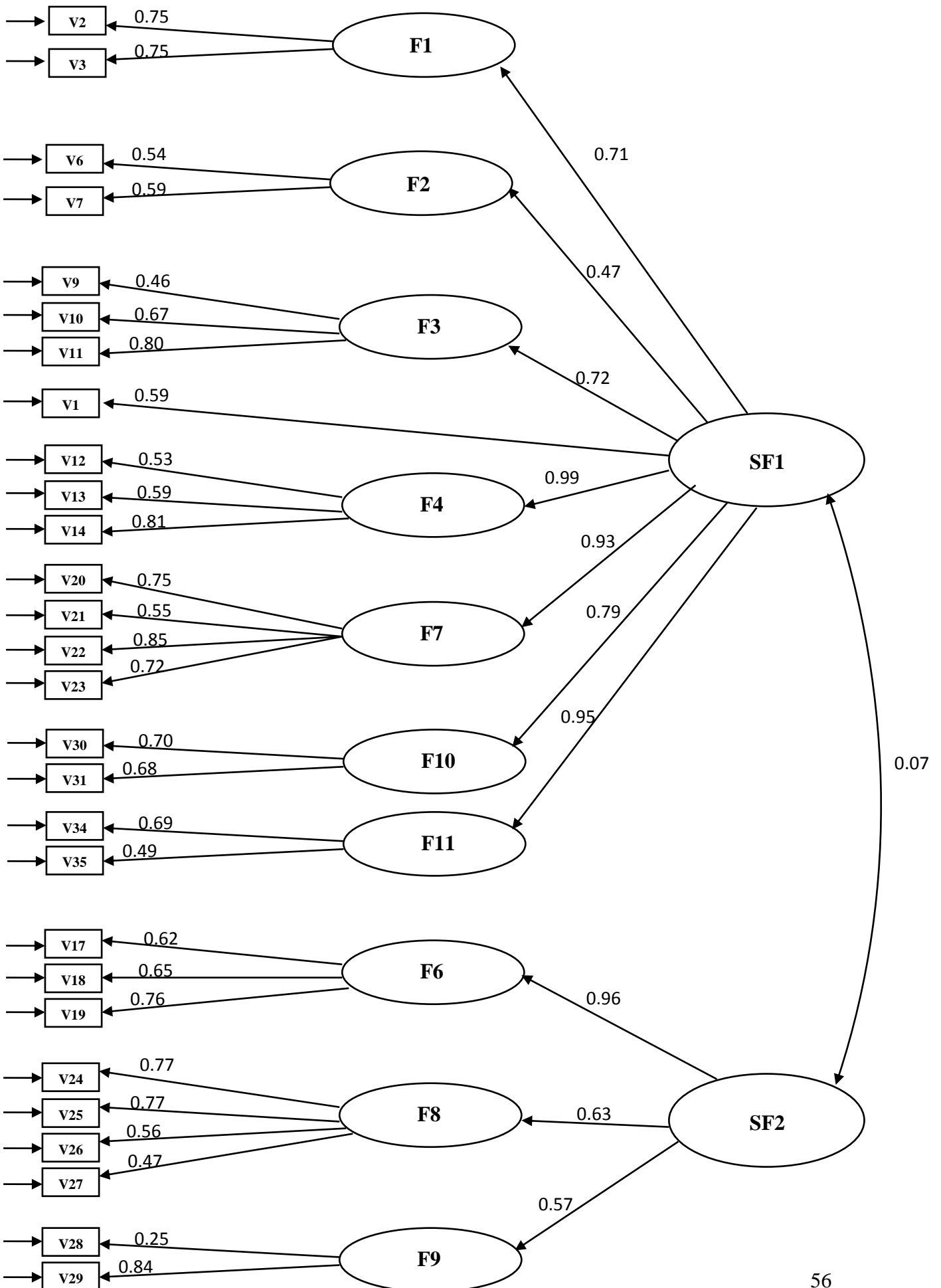




Figure 4: SEM results for Ireland

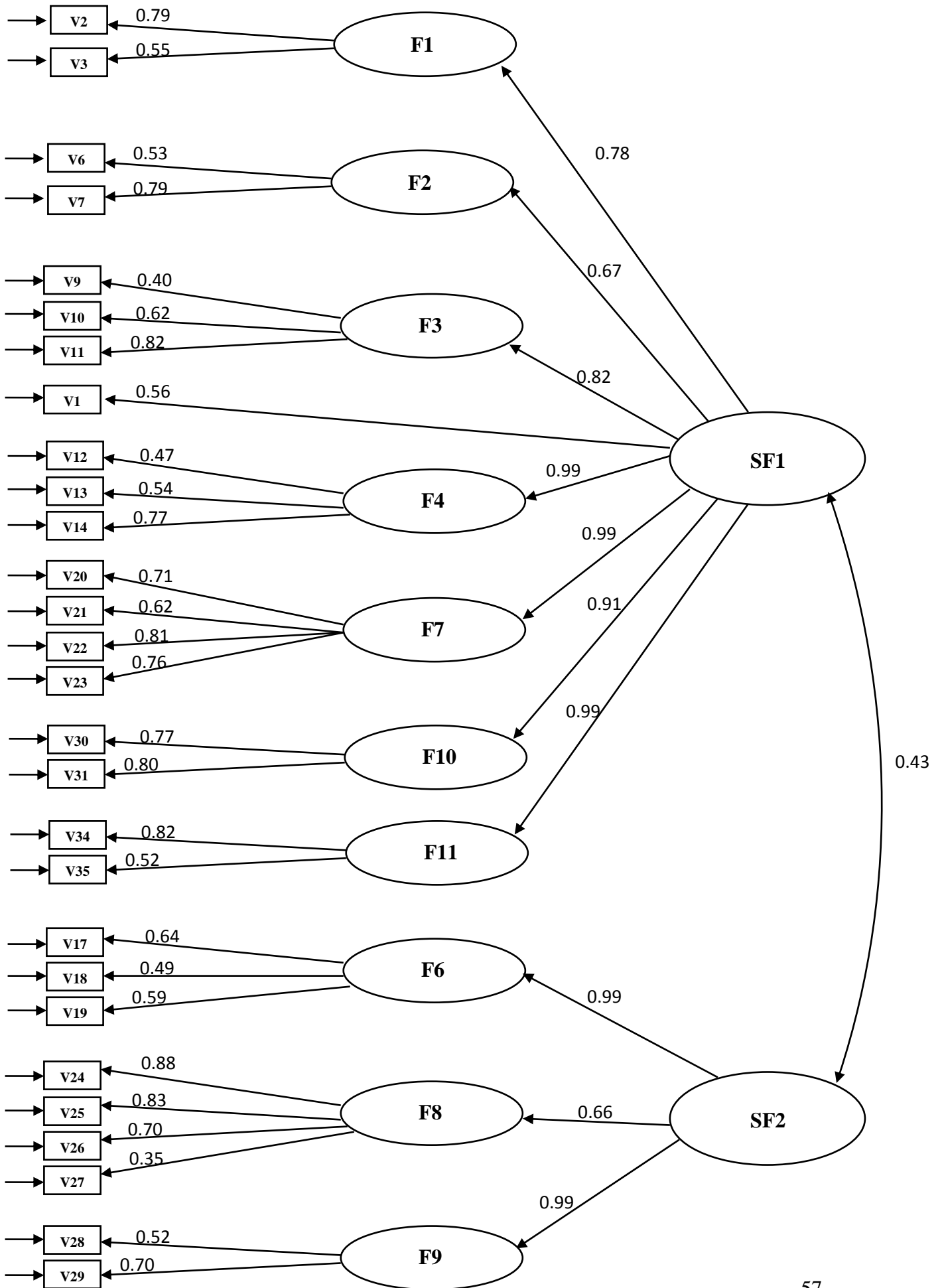


Figure 5: SEM results for Slovenia

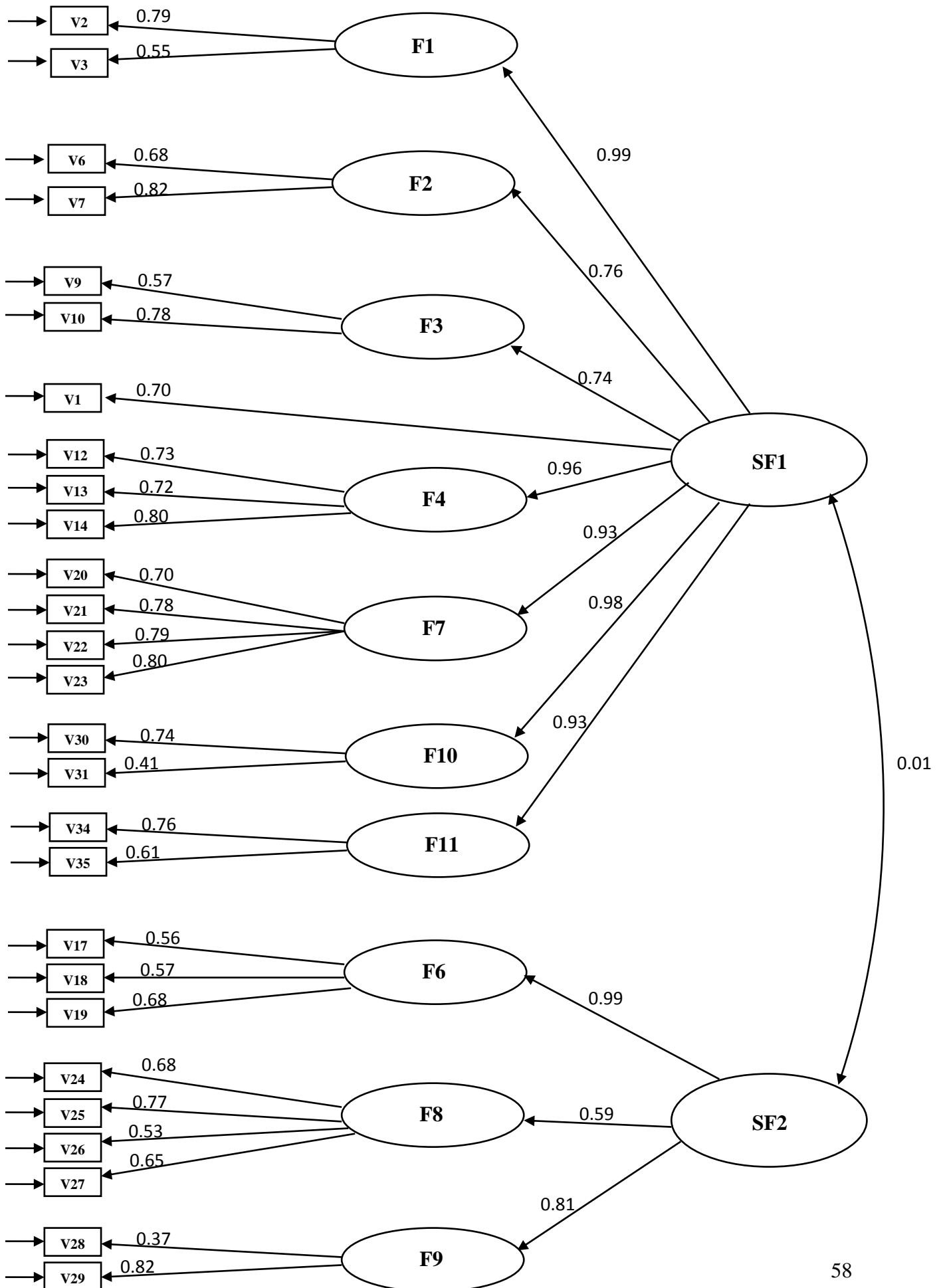
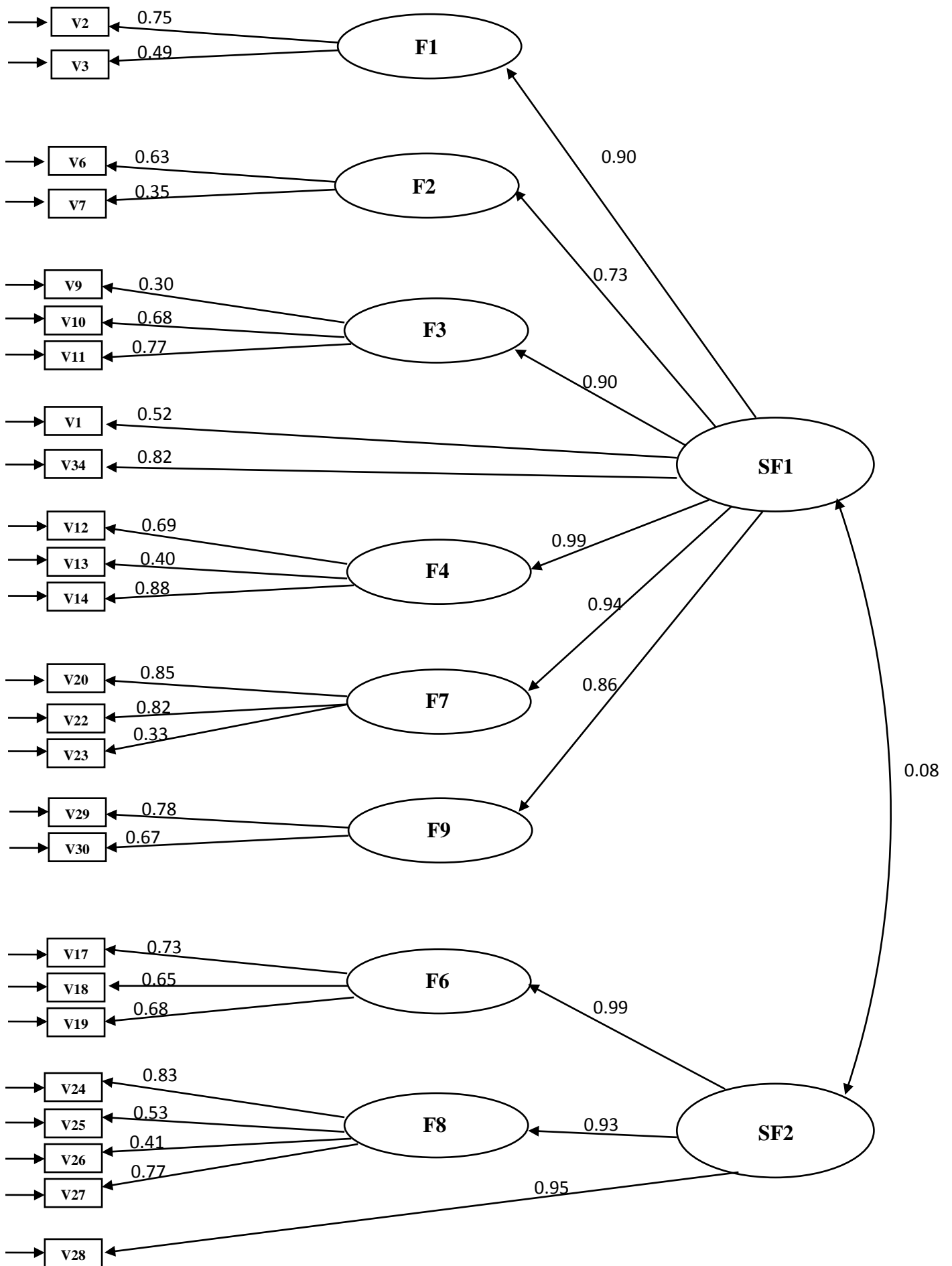


Figure 6: SEM results for Germany



As it was noted before, in Germany item 25 was not administered and therefore was excluded from the country analysis. Thus, from F9 (Questioning: Raising non-appropriate questions) only V28 (item 39) was included in the analysis. Also, items 9 (V33) from assessment and 17(V21) from classroom as a learning environment: teacher-student interaction were excluded from the analysis in Germany as their loadings were found to be low. The explanation of the figure for Germany is presented below:

**First Order Factors:**

**F1:** Modeling

**F2:** Structuring – Quantitative Characteristics

**F3:** Structuring – Qualitative Characteristics

**F4:** Application

**F6:** Management of Time

**F7:** Classroom as a Learning Environment – Qualitative characteristics: Teacher - Student interaction

**F8:** Classroom as a Learning Environment – Quantitative Characteristics: Dealing with Misbehaviour

**F9:** Questioning – Qualitative Characteristics

**V1:** Orientation

**V28:** Questioning: Raising non-appropriate questions

**V34:** Assessment

**Second Order Factors:**

**SF1:** Quality of Teaching

**SF2:** Quantity of Teaching (Management of Time, Misbehaviour & raising non – appropriate questions)