



BRILL

Journal of Empirical Theology 23 (2010) 159-178

Journal of
Empirical Theology



brill.nl/jet

Researching Confirmation Work in Europe: The Need for Multi-Level Analysis for Identifying Individual and Group Influences in Non-Formal Education

Wolfgang Ilg^a and Friedrich Schweitzer^b

a) Faculty of Protestant Theology, University of Tübingen, Germany
wolfgang.ilg@gmx.net

b) Faculty of Protestant Theology, University of Tübingen, Germany
friedrich.schweitzer@uni-tuebingen.de

Received 25 April 2010; accepted 12 July 2010

Abstract

This article includes a dual argument, about confirmation work as an object of research on the one hand, and about general methodological questions on the other. Based on an international comparative study on confirmation work with more than 28000 persons involved, the article presents the advantages of multi-level analysis (MLA) as a method for discerning effects on the levels of individuals and of groups. Discussing the methodological implications the article gives an exemplary introduction into MLA by presenting selected models referring to confirmands' satisfaction with their confirmation time in general and with church services in particular. It can be shown that 11-13% of the total variance is explained by factors on the group level. The group level factors include the duration of confirmation camp, the number of meetings and the aims of the workers as indicators for the educational concept of the team. Based on the intra-class correlation of MLA, the article suggests that the existence of fixed curricula in non-formal education settings like confirmation work enhances the uniformity of the programs and reduces the proportion of variance accounted for on the group level.

Keywords

confirmation work, youth work, Protestant Church, international comparative study, multi-level analysis (MLA), hierarchical linear modeling (HLM)

1. Introduction and Research Approach

With a first research project involving more than 28000 confirmands, parents and workers in seven European countries, the "International Network for Research and Development of Confirmation and Christian Youth Work" (www.confirmation-youthwork.eu) recently presented a study about one of the major educational settings within Protestant churches. Beyond the descriptive results (presented in Schweitzer, Ilg & Simojoki 2010), the study

allows for a deeper understanding of educational settings in general by discerning effects on the individual level, the group level and the country level. In building upon earlier publications from the project, the present paper offers more detailed interpretations based on multi-level analysis (MLA) or, as it is sometimes called, hierarchical linear modeling (HLM). The use of this method is needed whenever the research focuses on individuals who are parts of educational groups like confirmation groups or school classes and when sample sizes are large enough (typically several thousand respondents).

Until now, most research projects in the context of youth and religion have not used multi-level analyses. One recent example is the REDCo project (Religion in Education; Valk et al. 2009). The REDCo group administered questionnaires to 14 to 16 year old students in eight European countries asking them about their experiences with religion in school. In the REDCo study (as in many other international comparative studies), the authors compare the results on a country level. However, this way of analysis ignores the fact that also within a country the students are members of certain groups (e.g., school classes) that differ from each other. As long as the analysis draws on the dataset as a whole it assumes the same effects for all groups involved. Yet it is obvious that not only individual and country-related factors are influential but also group-specific factors exert an influence — and that the factors might interact across different levels.

Multi-level analysis (MLA) estimates the proportion of effects related to the different levels and can control for cross-level-interaction. The question to be answered in the following relates to such effects. What role do group effects really play? If it can be shown that such effects really exist and that they are not only of marginal influence, future research on religious pedagogies will have to include the multi-level approach much more than has been the case so far.

The aim of our paper is to provide an exemplary approach to multi-level analysis on the basis of the confirmation research data. We assume that most readers have not yet worked with MLA. Therefore we present a step by step introduction into the method and the need to use MLA based on the influence of group effects on the research results. The underlying question does not only refer to confirmation work but is a general one. In what respect do we profit from a multi-level view and how could this view be helpful for other studies?

2. The Shape of Confirmation Work

2.1 *How Confirmation Work is Organised*

In all of the seven countries involved — Germany (DE), Austria (AT), Switzerland (CH), Denmark (DK), Finland (FI), Norway (NO) and Sweden (SE) — confirmation work is set up by the Protestant churches as a one- or two-year course of study for adolescents aged 13 to 15. Recent changes in confirmation work have included the development from a purely catechetical setting towards a program with a wide variety of events and contents, which has received many impulses from youth work. More than half a million young people are confirmed in the seven countries every year. The participation rates in confirmation time (measured as the proportion of teenagers being confirmed in relation to all teenagers of the respective age group, regardless of their religious affiliation) range from 90% in Finland (which is even slightly more than the proportion of baptised Protestant teenagers) to 3% in predominantly Catholic Austria.

As Innanen & Krupka (2010) show, the duration of confirmation time varies between half a year (as it is the case in Finland and Sweden) and 20 months in northern and eastern Germany. Three countries (Germany, Finland and Norway) rely on some kind of curriculum, in the others (Austria, Switzerland, Denmark and Sweden), there exists only a rough definition of aims which leaves much space for the local parishes to organise their own programs. Sunday services are often called the “hidden curriculum” in confirmation work. Confirmands are usually obliged to attend a certain number of services, for example in Germany about 20 times a year. In most of the countries involved, camps have been established as an integral part of confirmation work. This is especially true for Finland, where the time spent at camp is by far the dominating experience of confirmation work.

2.2 *Previous Research*

An overview of previous studies in the field of confirmation work can be found in Schweitzer & Elsenbast (2009). In the light of the enormous participation in this field, it is surprising that earlier research was almost always restricted to regional studies within one country or parts of it. One exception was the attempt of the Lutheran World Foundation to bring together experiences on confirmation work from different countries by organizing consultations and conducting interviews with experts from different Lutheran churches all over the world (LWF 1995).

Studies with a quantitative approach are comparatively recent (Niemelä 2006/2008; Grahn, Eek & Pettersson 2007; Cramer, Ilg & Schweitzer 2009). Niemelä's continued t_1 - t_2 - t_3 study (Niemelä 2008) served as one of the starting points for the present study. By using regression analysis, she identifies which satisfaction ratings go along with a high overall satisfaction of confirmands. The general atmosphere in the community of young people and aspects related to the camp (accommodation, evening programs) are the most important factors here, while contents (the topics dealt with in the lessons) do not play an important role for Finnish confirmands. By interpreting correlations, Niemelä also identifies individual parameters with confirmands for predicting their satisfaction with confirmation training. She concludes from her data, that "a typical satisfied confirmand is a girl who right from the start has been positively disposed to confirmation training". Despite the differences in the satisfaction outcomes of the groups involved, her study restricts the search for the causes to individual factors of the confirmands.

The aim of the present paper is to go beyond this approach by identifying additional predictive factors on a group level. The strength of multi-level analysis is that effects on individual and on group level can be controlled simultaneously. A second advantage of MLA is its better estimation of the significance of results compared to traditional linear models. In regression models it is assumed, that error terms are unrelated. In nested data this assumption is often violated because of the shared experiences of individuals in a group, which leads to an alpha-level inflation (Kreft & De Leeuw 1998). The higher the homogeneity within a group, the more important it is to apply a MLA approach. This is why MLA always starts with the analysis of variance components accounted for on individual and on group level, which can be estimated by the intraclass correlation (ICC).

3. Method

3.1 Research Questions

The first two research questions refer to the satisfaction experiences of the confirmands using traditional methodology like descriptive statistics and regression analysis.

- a) *How satisfied are the confirmands with their experiences during confirmation time? Are there country-specific patterns of their satisfaction in the seven countries?*

- b) *Which satisfaction aspects show the strongest influence on the overall assessment?*

Following the Finnish regression analysis of Niemelä (2008) it can be expected that the feeling of community (“general atmosphere” in Niemelä’s study) and aspects connected to the camp experience have a high predictive power and that the satisfaction with the contents of the lessons is of less importance for predicting the confirmands’ overall assessment. Using our international data, we will test whether these findings can be generalised to other countries.

For the next two questions, we search for factors on the individual and on the group level for predicting overall satisfaction and the perception of church services.

- c) *Which predictors for the overall assessment can be identified on the individual and on the group level?*
d) *Which predictors for the confirmands’ perception of church services can be identified on the individual and on the group level?*

Niemelä found out that female gender and a positive disposition to church and Christian faith at the beginning of confirmation time make a significant contribution to the overall assessment. What additional effects can be found on the *group level*? For the pedagogical planning it is of special interest how the intensity of confirmation work and the camp experience relate to confirmands’ assessment of their confirmation time. Concerning the perception of church services, we are particularly interested in the question if the confirmands’ assessment of church services differs depending on the pedagogical program they experience, which is mirrored in the aims of the workers who lead the confirmation group.

The last question refers to the proportion of effects on the different levels in a comparative way between the countries.

- e) *How strongly do groups (here: their average degree of satisfaction) differ from each other within a country?*

One plausible hypothesis is that the differences are higher when there is a low level of standardisation, so that the four countries without obligatory curricula (Austria, Switzerland, Denmark and Sweden) should show a higher variety of results on the group level. This analysis could help in understanding to what extent the existence of a curriculum in non-formal educational settings determines the degree of uniformity in the group outcomes.

Table 1: Number of persons interviewed in the participating countries

country	TOTAL	DE	AT	CH ⁽¹⁾	DK	FI ⁽²⁾	NO	SE
units/groups	943	635	29	39	34	107	65	34
confirmands	19445	11513	540	598	1193	2176	2343	1082
workers	2386	1601	118	64	46	212	130	215
(parents)	6909	5788	240	246	–	–	635	–

⁽¹⁾ The Swiss results refer only to the cantonal Church of Zurich.

⁽²⁾ In Finland, the workers' survey did not address the 429 Young Confirmed Volunteers (YCVs) in the selected parishes. They received additional country-specific questionnaires.

3.2 Data

The international study on confirmation work was conducted as a t_1 - t_2 -study with questionnaires in the beginning of confirmation time (t_1 : autumn 2007) and shortly before confirmation (t_2 : spring/summer 2008). The participating groups were selected in order to gain a representative sample. The number of persons interviewed (Table 1) is roughly proportional to the number of confirmands in the countries. For example, Germany as the country with by far the most inhabitants and roughly 250000 confirmands per year accounts for 56% of all confirmands in the seven countries. Weighting factors were applied to balance regional distribution biases both within and between the countries. The questionnaires were translated into the respective languages German, Danish, Finnish, Norwegian and Swedish. English served as the communication basis for the development of the instruments.

Every participating parish received questionnaires for its confirmands and for all (employed and voluntary) workers who had been active at least three times during the actual confirmation period. In some countries, there were additional questionnaires for the parents and a t_0 -survey, conducted before the start of the study. In addition to the commonly used variables, each country also had the option to add more variables to the questionnaire. With few exceptions, this article refers solely to variables that were applied in all seven countries.

3.3 Research Instruments

Most items in the questionnaires were presented on a 7-point Likert-scale. The variables used as MLA-predictors have been z-standardised ($M=0$; $SD=1$) to facilitate the interpretation. As the proportion of missing values is low, no imputation has been applied. The following variables are used for the analyses in the present paper (extensive result tables are printed in Schweitzer, Ilg & Simojoki 2010):

- Satisfaction measures: In t_2 confirmands were asked “To what extent are you satisfied with...”, followed by an overall assessment (“the whole confirmation time”) and a list of 9 specific satisfaction categories (“minister[s]”, “camps”, etc.).
- Gender and age: 51.7% of the confirmands are female, 48.3% are male. The balance is almost even in all of the countries except Sweden, where female confirmands account for 60.1% of all confirmands. The age of the confirmands was asked in t_1 ($M_{\text{total}}=13.57$; $SD=.71$). Confirmands are youngest in Germany and Austria ($M<13.3$) and oldest in Switzerland, Sweden and Finland ($M>14.1$).
- Christian belief index (t_1): The closeness of the confirmands to the Christian faith when starting their confirmation time was measured by an index consisting of the 5 statements “God created the world”, “God loves all humans and cares about each one of us”, “Jesus has risen from the dead”, “Faith in God helps me in difficult situations” and “I believe in God”. The internal consistency is high throughout the countries, with an overall Cronbach Alpha of .87 (Germany .85, Austria .85, Switzerland .89, Denmark .83, Finland .91, Norway .88 and Sweden .87).
- Interest in church services: In the questionnaire at the beginning of their confirmation time (t_1) the confirmands were asked, how interested they were in certain topics. One item was “Course and meaning of Sunday services”. With an average of 3.53 ($SD=1.71$) on the 7-point scale, it ranked last among the 11 given topics.
- Information concerning the activity plan: This includes the overall number of meetings of the group (indicating the intensity of confirmation training) and the length of camps, measured by the number of nights. On average, there were $M=35.79$ meetings per group ($SD=19.79$), ranging from just under 14 in Austria and Finland to 43 in Germany (no information concerning this question is available in Sweden). The average number of nights in camp is $M=3.98$ ($SD=4.46$). Sweden ($M=11.08$) and Finland ($M=7.96$) have the highest number of camp nights. The result of Denmark ($M=.48$) shows that most parishes there do not use camps in confirmation work.
- In their t_1 -questionnaire the workers were asked to assess their personal aims for working with the confirmands. The present article uses two questions related to church services: “How important do you consider the topic ‘Course and meaning of Sunday services?’” ($M=5.52$; $SD=1.37$) and “During their confirmation time, the confirmands should experience forms of worship adequate for young people (e.g. youth church services)” ($M=5.82$; $SD=1.24$). These aims were aggregated on the group level, such reflecting the educational concept guiding the group processes.

3.4 Introduction to Multi-Level Analysis

In a nested structure (persons within groups), analyses without consideration of the multi-level structure can not do justice to certain effects, as Cronbach already pointed out in 1976: “The majority of studies of educational effects — whether classroom experiments, or evaluations of programs, or surveys — have collected and analyzed data in ways that conceal more than they reveal. The established methods have generated false conclusions in many studies”. The failure to distinguish between effects at the group level and individual effects leads to either ecological fallacy (inference from group statistics to individuals) or individualistic fallacy (inference from individual results to groups; Robinson 1950). When settings with a nested data structure are researched, neither the complete pooling of data (all individual data make up the dataset) nor the “no pooling” approach (mean values of each group make up the dataset) reflect the fact that individuals *and* groups are sources of variance. MLA, as it is common in group based studies like PISA, integrates aspects of both approaches and leads to “partial pooling”, using relative weights determined by the groups’ sample sizes (Gelman & Hill 2007; Raudenbush & Bryk 2002).

To provide a simple example, let us take two different confirmands’ groups, one with a high average satisfaction of the confirmands, the other with a low satisfaction. We assume that group 1 has taken part in a camp, whereas group 2 has not. The “complete pooling” approach would trace the causes for the satisfaction differences only on confirmands’ individual level and possibly conclude that the satisfaction of confirmands in t_2 is correlated to their religiosity in t_1 . The “no pooling” approach would leave individual aspects aside and simply compare the mean values of both groups. Thus, it might be concluded, that the camp experience results in higher satisfaction. Both ways of analysis are based on the expectancy of a fixed regression between a predictor X and the criteria variable Y . MLA adds new possibilities to this kind of analysis as it allows for different regressions for different groups and enables us to see the influences on the individual and on the group level at the same time. So it might be the case, that there is a regression between religiosity and satisfaction, but this regression varies systematically depending on the duration the group spent in camp. In the MLA approach the regression equation between religiosity (X) and satisfaction (Y) differs from group to group, as the intercept (β_0) and the slope (β_1) are not fixed but depend on group level variables like the number of camp nights (W). Equation 1 shows the basic idea of MLA. Each variable has two subscripts. The subscript i refers to the

individual, the subscript j refers to the groups, resulting in a different regression for every group.

Equation 1:

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + e_{ij}$$

intercept: $\beta_{0j} + \gamma_{00} + \gamma_{01} W_j + u_{0j}$

slope: $\beta_{1j} + \gamma_{10} + \gamma_{11} W_j + u_{1j}$

In the present article we allow for random intercepts, but use fixed slopes. Selected t_2 -variables of the confirmands are used as criteria variables. The predictors are taken from t_1 -variables of confirmands, structural information about the course of the confirmation time and workers' aims.

As a first step in every multi-level analysis, the intraclass correlation (ICC) is computed to investigate whether the partial pooling approach is promising. The ICC can be interpreted as the proportion of the total variance conveyed by the grouping. A high ICC indicates that groups are homogeneous and differ from each other, thus giving a strong argument for the need to apply MLA. If ICC is close to zero, the analysis on the individual level is sufficient and MLA does not provide additional insights.

4. Research Results

Research question a) was: *How satisfied are the confirmands with their experiences during confirmation time? Are there country-specific patterns of their satisfaction in the seven countries?*

Table 2 presents the mean values of the item section “To what extent are you satisfied with...” on a 7-point scale ranging from 1=“not satisfied at all” to 7=“totally satisfied”. The satisfaction of the confirmands with their experiences is high. For all the satisfaction items, the results for Finland and Sweden stand out. Confirmands in these countries are most satisfied with their experiences, especially with the camps. Confirmands in Switzerland and Denmark are most critical about their confirmation time. In five of the seven countries, camps are among the two favourite aspects (marked in bold); in four countries, ministers are; confirmands in three countries favour the “feeling of community”. Church services and prayers in the group are rated less satisfactory than all the other aspects — but they still have a mean score slightly above 4, the theoretical mean of the scale.

Table 2: Confirmands' satisfaction with different aspects (t₂): mean values

	SD	M	DE	AT	CH	DK	FI	NO	SE
	TOTAL	TOTAL							
<i>the whole confirmation time</i>	1.41	5.24	5.04	5.13	4.87	4.94	5.80	5.26	5.98
camp(s)	1.67	5.58	5.40	5.29	5.49	5.10	5.89	5.62	6.32
feeling of community	1.49	5.51	5.23	5.24	5.30	5.71	5.99	5.78	6.02
minister(s)	1.61	5.44	5.28	5.48	5.35	5.36	5.73	5.50	6.07
other teachers/workers	1.57	5.29	5.12	5.17	4.98	4.91	5.75	5.10	6.12
having fun	1.67	5.11	5.27	5.47	5.16	4.11	4.96	4.51	6.06
music, songs and singing	1.75	4.80	4.44	4.16	4.15	4.64	5.79	4.83	5.70
content/topics of lessons	1.39	4.75	4.42	4.33	4.53	5.01	5.53	4.79	5.34
prayers in the group	1.56	4.62	4.27	4.16	4.10	4.54	5.49	4.60	5.37
church services	1.61	4.49	4.29	4.32	4.17	4.25	5.25	4.52	4.96

The first line shows the overall assessment; the following items are ordered according to total results. Standard deviations are only given for the total. In each column, the two variables with the highest values are marked in bold type. If an item was not applicable, the confirmands could tick the box "we didn't have that". The number of valid answers therefore ranges from 15370 ("prayers in the group") to 17904 ("the whole confirmation time").

Research question b) was: *Which satisfaction aspects show the strongest influence on the overall assessment?*

Table 3 presents the regression coefficients for each country based on confirmands' satisfaction with the whole confirmation time as the dependent variable. The content/topics contribute highly to overall satisfaction in five of the countries; the "fun factor" is important in four countries. This is only different for Finland, where both variables are not among the favourites. Here the feeling of community and the satisfaction with camp show the highest regression weights. Thus, the factors presented by Niemelä (2008) are confirmed for Finland, but show a different pattern in the other countries.

The two aspects with the least satisfaction, "prayers in the group" and "church services", are not closely linked to the overall impression of the confirmands. In other words, although Sunday services are perceived as a somewhat boring compulsory exercise, this experience has only little impact on the overall positive assessment of confirmation time. It is also interesting that the confirmands' perception of the ministers is of much greater importance for their general impression than their feelings about other teachers and

Table 3: Factors explaining overall assessment (regression analysis): beta weights

	TOTAL	DE	AT	CH	DK	FI	NO	SE
<i>R</i> ²	.60	.57	.63	.62	.66	.60	.58	.66
content/topics of lessons	.27**	.25**	.15**	.24**	.39**	.15**	.26**	.25**
having fun	.20**	.24**	.25**	.18**	.21**	.11**	.04*	.35**
feeling of community	.16**	.12**	.12*	.10*	.09*	.31**	.20**	.13**
camp(s)	.13**	.15**	.19**	.20**	-.09*	.28**	.10**	.06*
minister(s)	.13**	.14**	.16**	.25**	.20**	.05*	.14**	.10**
music, songs and singing	.06**	.05**	-.02	.06	.00	.01	.06*	.06*
other teachers/workers	.04**	.04**	.14**	.01	.04	.03	.08**	.04
church services	.04**	.04**	.05	.01	-.02	.04	.10**	.02
prayers in the group	.03**	.03*	-.01	.04	.18**	.00	.05	-.01

** = significant at a 1% level; * = significant at a 5% level. In each column, the two variables with the highest values are marked in bold type.

workers. The two exceptions to this observation are Austria and Finland — in both countries, volunteers play a very important role; whereas for example in Germany, volunteers are represented numerically more than ministers but their involvement is most often limited to special events during the confirmation time. The results presented in Table 3 lead to the conclusion that contents and social aspects must both be seen as important factors for successful confirmation work. Alternatives, which were suggested earlier, and focus *either* on fun *or* topics are not adequate.

Regression analysis answers the question of which satisfaction factors contribute to the overall satisfaction but it cannot tell us which confirmands tend to give good or poor ratings and which group settings lead to high or low satisfaction. In the next step, relevant factors on an individual as well as on a group level concerning advantageous or disadvantageous conditions in the beginning of confirmation time (*t*₁) will be examined by using multi-level analysis.

The next research question c) “Which predictors for the overall assessment can be identified on the individual and on the group level?” requires the application of MLA.

The ICC model presented in Table 4 dissects the total variance in the proportions connected to individuals, groups and countries. 79.41% of the total variance can be explained by individual factors, 13.04% by group factors and 7.55% by country factors.

Table 4: Components of variance of the overall assessment of confirmands

	absolute	percent
within groups	1.5800	79.41%
between groups	.2594	13.04%
between countries	.1503	7.55%
total	1.9897	100.0%

Table 5: Components of variance of the overall assessment of confirmands: by countries

	DE	AT	CH	DK	FI	NO	SE
within groups	87.88%	88.94%	77.87%	78.56%	83.96%	86.79%	71.97%
between groups	12.12%	11.06%	22.13%	21.44%	16.04%	13.21%	28.03%
total	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Table 5 presents these proportions for each of the seven countries. The three countries with the highest proportions on group level are Sweden, Switzerland, and Denmark. The ICC test shows that a multi-level analysis appears to be promising, as relevant predictors for the overall assessment can be found both on the individual and on the group level.

For the benefit of readers who have not yet worked with multi-level analysis, we first describe the way of MLA notation. As an introductory example we start with a short initial MLA that only integrates three predictors on the individual level (and thus actually is a one-level analysis presented in a MLA notation). The criteria variable is the satisfaction with “the whole confirmation time”. The results in Table 6 can be interpreted in the following way. The estimate of 5.13 is the (theoretical) expected value for a confirmand with average Christian beliefs (t_1), an average age and male gender. The line “difference for girls” shows, that the expected value for girls is .14 higher than the one for boys. The stronger the confirmands agree with Christian beliefs in the beginning of confirmation time, the more positive their evaluation of their experience will be in the end. For a confirmand with a belief value one standard deviation above the mean, the expected value rises by .22. As the t-value for age is below the threshold of significance, the difference in the overall assessment explained by the age factor does not play a significant role.

Table 6: Multi-level analysis, model 1

Level	variable	estimate	Std. error	t-value
	(intercept)	5.13	.03	187.09**
I	difference for girls	.14	.02	6.58**
I	Christian beliefs in t_1 (z-value)	.22	.01	18.71**
I	age (z-value)	.03	.01	1.94

I = Individual Level.

For the next analyses, we omit the non-significant predictor “age” and we add predictors on the group level. In addition, we split the variable “Christian belief in t_1 ” in two components (each z-standardised). On the individual level, we use the value “belief (group mean centered)”, which is calculated as the deviation of one’s individual belief value from the mean belief value in the respective group. Thus, it shows the relative place of an individual compared to the average member of his/her group. On the group level, we use the mean value of each group, here called “average group belief”, indicating the group’s average value in comparison to all groups. Additionally, we include two variables describing group activities, the number of meetings during the confirmation year and the number of nights spent at a camp. The results in Table 7 suffer from the fact that the number of groups in the countries except from Germany is rather small for MLA. Still, in the total sample and in the case of Germany, interesting effects on the group level become apparent, and most of them show the same tendencies in the other countries, but fail to reach significance due to the small sample sizes. In addition to the two effects on the individual level (gender and belief), all three group level predictors have a significant impact. There is a slightly negative connection between the number of meetings during confirmation time and the satisfaction of confirmands. Clear evidence can be seen for the positive effect of camps. The length of the camp period correlates significantly with confirmands’ satisfaction in four of the countries. The high negative effect in Denmark is not significant and probably is an artefact due to the fact that only 9 parishes in the Danish sample went to camp at all.

The significant predictive power of gender and Christian belief on individual level replicate the results of Niemelä (2008). The longer the camp period lasts, the more positive is the confirmands’ perception of confirmation time — an effect which is significant in four of the seven countries. There is some evidence for a reverse effect of the number of meetings and overall assessment, but the effect is small and differs throughout the countries.

Table 7: Multi-level analysis, model 2

Level	variable	TOTAL	DE	AT	CH	DK	FI	NO	SE
	(intercept)	5.12**	4.98**	5.30**	5.17**	4.27	5.66**	5.85**	6.14**
I	difference for girls	.14**	.13**	.13	-.06	.17	.14*	.16*	.24**
I	belief (group mean centered, z-value)	.20**	.20**	.33**	.08	.33**	.15**	.26**	.18**
G	average group belief (z-value)	.09**	.16**	.22	.31	.36	.15	.05	.46
G	number of meetings (z-value)	-.11**	-.02	.09	.10	-.20	.01	.13	⁽¹⁾
G	number of nights in camp (z-value)	.32**	.19**	1.07**	.03	-.96	.28	.93*	.15*

I = Individual Level; G = Group Level; ⁽¹⁾ In Sweden, this question was not asked.

Table 8: Components of variance of the assessment of church services by confirmands per country

	TOTAL	DE	AT	CH	DK	FI	NO	SE
within groups	83.10%	89.54%	82.92%	84.98%	84.39%	86.03%	90.23%	77.63%
between groups	11.42%	10.46%	17.08%	15.02%	15.61%	13.97%	9.77%	22.37%
between countries	5.49%							
total	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Research question d) asks: *Which predictors for the confirmands' perception of church services can be identified on the individual and on the group level?*

The dependent variable for the following MLA is satisfaction with church services (t_2). Table 8 presents the ICC for the total dataset and each of the seven countries. 13% of the variance is caused by the difference between groups, nearly 8% by the difference between countries. Again, Sweden, Denmark and Switzerland — and this time, also Austria — have higher proportions between groups than the other countries.

For the MLA we use the same individual predictors as above. On the individual level we add the degree of interest in learning about church services as stated by the confirmands in t_1 . On the group level we introduce a new type of variable: workers' aims. Assuming that not only the events (like number of meetings, number of camp nights, etc.) shape the effects of confirmation

time but also the intentions of the responsible workers, we add indicators for these pedagogical aims to the MLA. We use the aggregated workers' aims concerning learning about the meaning of church services and concerning the motivation to organise youth-adequate services.

The multi-level analysis presented in Table 9 shows clear results throughout all the countries with respect to two effects on the individual level. Church services are rated more favourably by confirmands who have been interested in the course and meaning of Sunday services from the beginning of their confirmation time, and who agree more strongly with statements concerning the Christian faith compared to the average confirmands of their group. Girls tend to like services more than boys, but the effect is not as clear as for the overall assessment. The age factor is significant on the total level, but the effects within the countries show a rather unclear picture. The average religiosity of the group is a supporting factor in Germany, Norway and especially in Sweden. Concerning the workers' aims, there is a positive effect concerning the aim to organise youth-adequate services (significant only for the total dataset), but no significant (rather tending to a negative) effect concerning the aim to teach course and meaning of services.

These results confirm the hypothesis that the confirmands' attitude towards church services is strongly dependent on their religiosity when starting confirmation time and on the degree of their personal interest in learning about services. The average group belief is an important predictive factor in Sweden, which probably has to do with the option for young Swedes to choose between different types of confirmation work — thus it can be assumed that highly church-affiliated confirmands prefer certain groups, for which the average group belief is a telling indicator. The predictive power of the workers' aims can be found, but it is less than expected. Nevertheless, the results concerning the aims of the workers lead to an interesting interpretation. The aim to provide confirmands with an intellectual understanding of church services does not account for a rise in the young people's assessment — there is rather a tendency towards the opposite. In contrast, a positive effect can be stated for the aim to organise youth-adequate services during confirmation time. This is in line with the finding that confirmands appreciate services much more if they experience youth-adequate services during their confirmation time and can themselves participate in preparing services for the congregation (Ilg, Schweitzer & Elsenbast 2009, 143).

Table 9: Multi-level analysis for assessment of church services

Level	variable	TOTAL	DE	AT	CH	DK	FI	NO	SE
	(intercept)	4.38**	4.13**	4.49**	4.16**	4.45**	5.05**	4.79**	5.33**
I	difference for girls	.11**	.07*	-.24	-.07	.03	.28**	.11	.15
I	age (z-value)	.07**	-.01	.02	.04	-.12	.14	-.01	.15
I	confirmand's interest in services t_1 (z-value)	.35**	.36**	.21*	.25**	.49**	.15**	.34**	.40**
I	belief (group mean centered, z-value)	.21**	.19**	.22*	.14*	.33**	.24**	.18**	.17**
G	average group belief (z-value)	.05	.18**	-.16	-.05	.13	.11	.19*	.75*
G	workers' aim: course and meaning of services (z-value)	-.02	-.00	-.04	-.10	.12	.05	.14	-.34
G	workers' aim: organise youth-adequate services (z-value)	.08**	.04	.28	-.03	.03	.16	.01	.26

I = Individual Level; G= Group Level

The last research question e) was: *How strongly do groups (here: their average degree of satisfaction) differ from each other within a country?*

For answering this question we draw on the intraclass correlations (ICC) for both dependent variables: overall assessment and satisfaction with church services. The intraclass correlations (Table 5 and Table 8) reveal the greatest satisfaction differences between the groups for Sweden, Switzerland, Denmark and (only for church service assessment) Austria. Those are the four countries without a general curriculum. Obviously, the lack of common regulations for non-formal educational settings enhances the variety between the groups. One might assume that this lack of common guidelines leads to unfavourable results — and indeed Denmark and Switzerland have the least satisfied confirmands concerning overall assessment, whereas Austria ranks in the middle of the countries (Table 2). But the result of Sweden with the most satisfied confirmands of all seven countries contradicts this hypothesis. The lack of a national curriculum apparently leads to different results between the groups but not necessarily to bad ones.

5. Summary and Discussion

5.1 *Summary of the Results*

The satisfaction of confirmands with their confirmation time is high, especially concerning the camps. The results of the regression analysis show the importance of differentiating the analyses per country. Only the Finnish results in our international study are in line with Niemelä's findings from a former Finnish study, where community and the camp experience dominate over all other aspects concerning their role for the overall assessment. In the other countries, it is the combination of fun and contents that accounts for a good or bad overall impression of the confirmation time.

While the presented regression analysis shows which satisfaction factors are strongly correlated with the overall satisfaction, it does not help in *predicting* which personal or group related factors account for higher satisfaction rates. For example, we know from regression analysis that confirmands who are highly satisfied with their camp, tend to rate their whole confirmation time more positively. But this analysis does not provide information about what kind of camp or which pedagogical settings are helpful for the positive experience. Here multi-level analysis integrates individual factors (especially gender and Christian belief) as well as group level predictors (duration of camp, number of meetings during the whole confirmation time) to predict the overall satisfaction in respect to both levels.

Concerning the group level predictors, we find evidence for the hypothesis that the duration of camps accounts for a higher satisfaction with confirmation time (even after controlling for individual factors), whereas a high number of meetings during the whole confirmation time tends to predict a more negative overall assessment. With regard to the church services, the workers' aim to organise youth-adequate services advances the satisfaction of the confirmands but the aim to teach about the course and meaning of church services does not show this positive effect.

The differences between the groups are highest for the countries without a general curriculum. In other words, where there is a general curriculum, the results of different groups show a higher similarity and the within-group homogeneity (mirrored in the ICC results) is smaller.

5.2 *The Advantages of a Multi-Level Analysis Approach*

The two analyses for overall assessment and satisfaction with church services give an impression of the usefulness of multi-level analysis. This method allows to discern effects on micro-, meso- and macro-level. Although the

individual differences account for the highest proportion of variance, a significant proportion can also be found on the level of groups and countries. Thus MLA enhances the proportion of explained variance and can help to identify activities and pedagogical aims that are of importance for pedagogical group settings. Especially the idea of operationalising the agenda of a team by asking for the pedagogic aims of the workers seems to be a promising approach that grasps the real aims of a local team much more precisely than simply relying on general descriptions like the analysis of a curriculum. Moreover, in a more general sense, the present analysis proves that influences at the group level do play a role. Not taking them into account would distort our understanding of the data. In the present context, their proportional influence may appear limited when compared to the individual influences. Nevertheless they cannot be neglected. Instead, given their existence, it must be concluded that future studies should definitely offer a control for group level influences whenever applicable. Otherwise incorrect conclusions can hardly be avoided.

We therefore suggest that MLA should be taken into account in more studies in the field of (religious) pedagogy. Whenever data is collected in a nested design and when a sufficient data pool is given, multi-level analysis should complement the classical uni-level regression analyses. Especially in the context of non-formal education this has rarely been done so far. One important reason is the fact that MLA requires a large dataset with hundreds of groups involved as it often is the case in school settings (like in the PISA studies). It is challenging but not impossible to gain such large scale datasets also in selected non-formal educational settings like youth work or international youth encounters (cf. Dubiski/Ilg 2008).

5.3 *Confirmation as a Non-Formal Educational Setting*

Finally, we ask what light our results shed on educational settings in general. Today's confirmation work has become one of the most visible and well-structured examples of non-formal education with clear contributions to civil society. Although the importance of non-formal education, i.e., forms of institutionalised educational programs other than compulsory education, has recently been emphasised (Cedefop 2009; Rauschenbach 2009), there are few studies that scrutinize the relevant group effects. Our international study on confirmation work gives interesting insights into this kind of non-formal educational settings. One important aspect is the question if a national curriculum has an effect on the homogeneity of confirmation work. As the ICC analyses show, countries with a general curriculum show less differences between the groups. A curricular framework thus can be seen as a helpful

means for providing a basic level of standardisation in non-formal educational settings.

Camps receive the best ratings of all aspects that take place in confirmation time. The MLA of confirmands' overall assessment indicates that the length of camp improves confirmands' satisfaction with what they perceive during confirmation time. Therefore camps can be seen as a key factor for the future success of confirmation work — resulting in the recommendation that countries without a camp tradition like Denmark should be encouraged to introduce these intensive group outings. Despite the fact, that the youth work-like “fun aspects” account for much of the success of confirmation work, it must not be overlooked — as the regression analysis revealed — that the content and topics of lessons are even more important for confirmands' overall assessment. The secret of successful confirmation work — and probably also for similar efforts in other settings of non-formal education — seems to lie in the combination of fun and content.

The results of our study lead to several new research questions. Is the existence of set curricula in non-formal educational settings generally associated with a greater uniformity in the results of these pedagogical activities? Can the same type of relation between workers' aims and group outcomes be found for other areas of aims and in other non-formal educational settings? We suggest that multi-level analysis should be used more often in the future for answering these kinds of questions, especially in the context of non-formal education.

References

- Cedefop — European Centre for the Development of Vocational Training (2009). *European guidelines for validating non-formal and informal learning*. Luxembourg: Office for Official Publications of the European Communities.
- Cramer, C., W. Ilg & F. Schweitzer (2009). *Reform von Konfirmandenarbeit — wissenschaftlich begleitet. Eine Studie in der Evangelischen Landeskirche in Württemberg* [Reforms in confirmation work — a study in the Protestant Church of Württemberg]. Gütersloh: Gütersloher Verlagshaus.
- Cronbach, L.J. (1976). *Research on Classrooms and Schools: Formulation of Questions, Design and Analysis*. Stanford Evaluation Consortium.
- Dubiski, J. & W. Ilg (Eds.) (2008). *Evaluation Internationaler Jugendbegegnungen. Ein Verfahren zur Auswertung von Begegnungen* [Evaluation of international youth encounters. Tools for a self-evaluation]. Berlin / Paris / Warschau: Deutsch-Französisches Jugendwerk / Deutsch Polnisches Jugendwerk. [also available in Polish (2009) and French (2010)]
- Gelman, A. & J. Hill (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.
- Grahn, N., J. Eek & P. Petterson (2007). *Vägar framåt i Svenska kyrkans konfirmandarbete*. [The way of confirmation work in the Swedish church] Karlstad: Karlstads stift.

- Ilg, W., F. Schweitzer & V. Elsenbast (2009). *Konfirmandenarbeit in Deutschland. Empirische Einblicke — Herausforderungen — Perspektiven* [Confirmation work in Germany. Empirical results — challenges — perspectives]. Gütersloh: Gütersloher Verlagshaus.
- Innanen, T. & B. Krupka (2010). Planning and Teaching in Confirmation Work. In: Schweitzer, F., W. Ilg & H. Simojoki (Eds.) (2010). *Confirmation Work in Europe: Empirical Results, Experiences and Challenges. A Comparative Study from Seven Countries* (pp. 222-234). Gütersloh: Gütersloher Verlagshaus.
- Kreft, I.G.G. & J. De Leeuw (1998). *Introducing multilevel modelling*. London: Sage Publications.
- LWF — Lutheran World Federation (1995). *Studie über Konfirmandenarbeit. Berichte aus der ganzen Welt* [Study on confirmation work. Reports from the whole world]. Geneva: LWF.
- Niemelä, K. (2006). The quality and effectiveness of confirmation classes in Finland. *Journal of Beliefs and Values*, 27, 177-190.
- . (2008). *Does confirmation training really matter? A longitudinal study of the quality and effectiveness of confirmation training in Finland*. Tampere: Church Research Institute.
- Raudenbush, St. W. & A.S. Bryk (2002). *Hierarchical Linear Models: applications and data analysis methods*. Thousand Oaks/CA: Sage.
- Rauschenbach, Th. (2009). *Zukunftschance Bildung. Familie, Jugendhilfe und Schule in neuer Allianz*. [Education for the future. Towards a new alliance of family, youth welfare and school]. Weinheim/München: Juventa.
- Robinson, W.S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review* 15, 351–357.
- Schweitzer, F. & V. Elsenbast (Eds.) (2009). *Konfirmandenarbeit erforschen. Ziele — Erfahrungen — Perspektiven* [Researching confirmation work: Aims — experiences — perspectives]. Gütersloh: Gütersloher Verlagshaus.
- Schweitzer, F., W. Ilg & H. Simojoki (Eds.) (2010). *Confirmation Work in Europe: Empirical Results, Experiences and Challenges. A Comparative Study from Seven Countries*. Gütersloh: Gütersloher Verlagshaus.
- Valk, P., G. Betram-Troost, M. Friederici, C. Béraud (2009). *Teenagers' Perspectives on the Role of Religion in their Lives, Schools and Societies. A European Quantitative Study*. Münster: Waxmann.