www.ijres.net

International Journal of Research in Education and Science (IJRES)

Findings from TIMSS 2007: What Drives Utilization of Inquiry-based Science Instruction?

Aliya Kuzhabekova
Nazarbayev University, Kazakhstan,
aliya.kuzhabekova@nu.edu.kz

To cite this article:

Kuzhabekova, A. (2015). Findings from TIMSS 2007: What drives utilization of inquiry-based science instruction? *International Journal of Research in Education and Science (IJRES)*, *1*(2), 142-150.

International Journal of Research in Education and Science
Volume 1, Issue 2, Summer 2015 ISSN: 2148-9955

Findings from TIMSS 2007: What Drives Utilization of Inquiry-Based Science Instruction?

Aliya Kuzhabekova*
Nazarbayev University, Kazakhstan

Abstract

Prior research has shown that greatest student achievement in sciences is attributed to inquiry-based instructional approach, in which the goal of science teaching is nurturing attitudes and skills necessary for independent quest for scientific knowledge. While prior research has clearly demonstrated positive instructional effects of inquiry-based approach, there is little understanding of what factors determine utilization of the instructional methodology in classrooms. This study uses hierarchal-linear statistical analysis to explore the effect of some teacher, school and country-level factors that might determine utilization of inquiry-based approach around the world. Country level data for the study came from RAND's ranking of countries in terms of their science and technology capacity and the GDP data by the International Monetary Fund. Teacher and school level data were obtained from the international database for Trends in International Mathematics and Science Study 2007. Based on the results of intraclass correlation analysis, the study revealed that most of the variability is determined at the school level and, in addition, some slight variability is explained by differences among countries. The exploratory analysis was able to identify potential predictors at the individual level. None of the explored predictors at the school and country level were significant.

Keywords: TIMSS 2007; Inquiry-based science instruction

Introduction

Prior research has shown that greatest student achievement in sciences is attributed to the so called *inquiry-based instructional approach*. Inquiry-oriented science instruction has been characterized in a variety of ways over the years (Collins, 1986; DeBoer, 1991; National Science Board, 1991; Rakow, 1986). In summary, in this approach the goal of science teaching is not mere transfer of knowledge, but rather nurturing attitudes and skills necessary for independent quest for scientific knowledge. Procedurally, the process of inquiry-based teaching should resemble as much as possible the actual process of scientific discovery. Students in inquiry-based classrooms act as mini-scientists, formulating real-life problems that need to be resolved via scientific exploration, involving in observation of natural phenomena or experimentation to collect relevant data, analyzing the data, drawing and communicating conclusions with minimal supervision from the teacher and, frequently, collaborating with other students in class.

Substantial evidence has been generated to show the effectiveness of inquiry-based teaching. Minner, Levy and Century (2010) implemented a meta-analysis of 138 studies on inquiry-based learning published over the period of 1984-2002. They concluded that most of the studies indicate a clear positive trend in favoring inquiry-based instruction. Over half of the analyzed studies found a positive impact of inquiry-based instruction on students' content learning and retention. Inquiry-based approach has also been found beneficial in understanding of science processes (Lindberg, 1990), vocabulary knowledge and conceptual understanding (Gautreau & Binns, 2012; Lloyd & Contreras, 1985), critical thinking (Narode et al., 1987), positive attitudes toward science (Kyle, 1985; Pai-Lu Wu et al., 2014; Rakow, 1986; Sandoval & Harven, 2011), higher achievement on tests of procedural knowledge (Glasson, 1989; Hung, 2009), and construction of logico-mathematical knowledge (Staver, 1986). Some recent studies found that inquiry based instruction is beneficial only if properly guided by teachers (Alfieri et al., 2011; Barthlow & Watson, 2014; Furtak et al., 2012).

While much research has been conducted on the instructional effects of inquiry-based approach, there is little understanding of what factors determine utilization of the instructional methodology in classrooms. Meanwhile, understanding of potential enablers and barriers to implementation of inquiry-based approach could increase the effectiveness and efficiency of policy interventions aimed at improving science instruction. Few studies (Garcia,

* Corresponding Author: *Aliya Kuzhabekova, aliya.kuzhabekova@nu.edu.kz*

2003; Inda, 2013) attempted to explore the relationship between teacher characteristics (experience, background in science, subject matter knowledge, and skills in and attitudes towards inquiry based instruction). Only one study by Pea (2012) tried to explore contextual factors, which may influence the ability of teachers to use inquiry based approaches. However the study measured teachers' subjective perceptions about the importance of the factors rather than attempted to explore the relationship between contextual characteristics and actual utilization of inquiry based methods by teachers. This study uses hierarchal-linear statistical analysis (Raudenbush & Bryk, 2002) of recent data from the Trends in International Mathematics and Science Study 2007 to explore the effect of some teacher, school and country-level variables that might determine actual utilization of inquiry-based approach in science instruction around the world.

Data

The data for the study came from three sources. Country level data came from a ranking of countries in terms of their science and technology capacity, which was developed by RAND (Wagner, Brahmakulam, Jackson, Wong, & Yoda, 2001), and the GDP data by the International Monetary Fund (2010). Teacher and school level data were obtained from the international database for Trends in International Mathematics and Science Study, which was conducted by the International Association for Evaluation of Educational Achievement (IEA) in 2007. The study utilized a complex sampling procedure and collected data using achievement tests and a variety of accompanying background surveys, the details of which can be obtained from the IEA website (IEA, 2007). The database comprises student achievement data as well as student, teacher, school, and curricular background data for 59 countries and 8 benchmarking participants (IEA, 2007). For the purposes of this study, only the data from eight-grade science teacher- and school- questionnaires was utilized and the student achievement and student-questionnaire data was disregarded. The data from benchmarking participants was also excluded from the analysis.

One of the unavoidable properties of the large scale databases that create complications for any statistical analysis is the missing data problem. International TIMSS database is especially notorious for the problem due to the technical difficulty in organizing data collection at the international scale. Prior studies have implemented missing data analysis, including the analysis of missing data in background survey items, and found that the data in TIMSS 2007 is missing at random (Wiberg & Andersson, 2010). MAR data has been shown to result in biased, lest efficient parameter estimates and lower statistical power (Azen, Van Guilder & Hill, 1989; Haitovksy, 1968; Kim & Curry, 1977). IEA uses EM imputation (Dempster, Laird & Rubin, 1977) to deal with missing data in the achievement tests and leaves handling of missing background data to the discretion of researchers. Prior studies used imputation for missing background data (Wiberg & Anderson, 2010). However they either analyzed a subset of countries or employed a two level model. In this study, due to the amount of data necessary for three-level analysis data imputation was practically unfeasible and leastwise deletion was utilized. Overall, the study was based on the data from 264, 527 teachers in 5,279 schools from 40 countries.

Model

The dependent variable of interest was a composite variable designed from the teacher questionnaire responses, which were viewed as most relevant for assessment of the extent of utilization of inquiry based method. The composite variable was calculated as the average score of a teacher on the items assessing the frequency with which the teacher asks students to: (1) design or plan experiments or investigations, (2) conduct experiments or investigations, (3) work together in small groups to conduct experiment or investigations (item 17 in Science Teacher Questionnaire).

The analysis was conducted in two steps. First, an unconditional one-way random effects ANOVA was fitted to learn if there is significant between group variability to necessitate multilevel modeling. Second, the multilevel model was run to represent variation at the teacher, school and country levels.

Teacher level (Level I) model was specified as follows:

$$Yijk = \pi 0jk + \pi 1jk*(AGEijk-\overline{AGE}) + \pi 2jk*(SEXijk - \overline{SEX}) + \pi 3jk*(EXPERijk - \overline{EXPER}) + \pi 4jk*(EDUijk-\overline{EDU}) + \pi 5jk*(PERCEPijk - \overline{PERCEP}) + \pi 6jk*(CL\_SIZEijk - \overline{CL\_SIZE}) + \varepsilon ijk ,$$

where Yijk is the extent of utilization of inquiry-based instruction by a teacher i in school j in country k; AGE, SEX, EXPER, EDU, PERCEP, CL_SIZE are independent variables representing teacher's age, sex, level of completed education, perception of student desire to do well in school (average for the class of the student), and class size correspondingly with subscripts i, j, k defining the values of these variables for individual within a school and a country correspondingly; π0jk is Level I intercept representing the grand mean frequency of utilization of the inquiry-based method in country k; π1jk, π2jk, π3jk, π4jk, π5jk, π6jk – represent LEVEL I slopes indicating how much each of the independent variables contributes to variation in the dependent variable beyond the grand mean holding all other variables constant; and εijk is the random effect associated with teacher i in school j and country k. Note that in the model all independent variables are grand mean centered to improve interpretability of the intercept. Grand mean centering was chosen over group mean centering because this study is concerned with estimating effects at all three levels; and in grand mean centering higher-level effects are adjusted for Level I effects, thought at the expense of potential misspecification of Level I coefficients in case of a higher-level model misspecification (Raudenbush & Bryk, 2002).

School-level (Level II) model was specified as follows:

$$\pi 0jk = \beta 00k + \beta 01k*\%LOW\_SESjk + \beta 02k*SCH\_SIZEjk + \beta 03k*PROF\_DEV\_CONTENTjk + \beta 04k*PROF\_DEV\_SKILSjk + \beta 05k*INSTR\_MAT\_SHORTAGEjk + \beta 06k*LAB\_EQUIP\_SHORTAGEjk + \beta 07k*TEACHER\_SHORTAGEjk + r0jk;$$

$$\pi 1jk = \beta 10k;\ \pi 2jk = \beta 20k;\ \pi 3jk = \beta 30k;\ \pi 4jk = \beta 40k;\ \pi 5jk = \beta 50k;\ \pi 6jk = \beta 60k;$$

where the predicted variables are the intercepts (π0jk) and slopes (π1jk , π2jk, π3jk, π4jk, π5jk, π6jk) from Level I model representing grand mean frequency of use of inquiry-based instruction and school-specific effects of Level I variables on the dependent variable; β00k is average intercept for country k, also representing mean frequency of use of inquiry-based techniques across all schools in a country; %LOW_SES, SCH_SIZE, PDEV_CONTENT, PDEV_SKILLS, MAT_SHORT, EQUIP_SHORT, TEACHER_SHORT are school-level predictor variables for prediction of Level I intercept representing percent of low SES students in a school, school size, opportunities for professional development in the subject matter, opportunities for professional development in pedagogy, extent of shortage of instructional materials, laboratory equipment and teachers respectively; β01k, β02k, β03k, β04k, β05k, β06k, β07k – are Level II slopes, indicating the magnitude and direction of influence of each of the corresponding predictor variables (fixed effects) on the Level I intercept holding the effect of other variables constant; β10k, β20k, β30k, β40k, β50k, β60k- are Level II fixed effects associated with the dependent variables at Level I . Note that at Level II only intercepts are random, while slopes are constrained to be the same across countries. This constraint was imposed to reduce the number of parameters to be estimated and was viewed as appropriate because the study's primary concern is the effect of the predictors on the mean frequency of utilization of inquiry-based methods (intercepts).

Country level (Level III) model was specified as follows:

$$\beta 00k = \gamma 000 + \gamma 001GDPk + \gamma 002STI\_RANKk + u00k ;$$

$$\beta 01k= \gamma 010;\ \beta 02k=\gamma 020;\ \beta 03k= \gamma 030;\ \beta 04k=\gamma 040;\ \beta 05k= \gamma 050;\ \beta 06k=\gamma 060;\ \beta 07k= \gamma 070;\ \beta 10k=\gamma 100;\ \beta 20k= \gamma 200;\ \beta 30k=\gamma 300;\ \beta 400k= \gamma 400;\ \beta 50k=\gamma 500;\ \beta 60k=\gamma 600;$$

where dependent variables are intercepts (β00k) from Level II, representing average intercept (mean frequency of use of inquiry based methods across schools in a country), and Level II slopes (β01k; β02k; β03k; β04k; β05k; β06k; β07k; β10k; β20k; β30k; β400k; β50k; β60k) representing average intercept (mean effect of school level variables) for a country k; *γ000* – is Level III intercept indicating grand mean frequency of use of inquiry based instruction across all countries in the world; GDP and STI_RANK – are Level III predictor variables, indicating GDP per capita and science and technology capacity rank of a country respectively with subscript k indicating the value of the variables for a specific country; γ001 and γ002 – are Level III slopes representing the direction and magnitude of the effect of a country' GDP and STI_RANK on the mean frequency of use of inquiry –based methods across schools in the country; u00k – is random country level error; and γ100; γ200; γ300; γ400; γ500; γ600 – are intercepts, representing fixed effects of Level II variables. Note that at Level III only intercepts are random, while slopes are constrained to be the same across countries[*]. As in the case of Level II model, the constraint was imposed to reduce the number of parameters to be estimated.

[*] Specification of the complete mixed model is omitted from the paper because of space limitations and is left as exercise to the reader.

Results

Fitting the unconditional model indicated that multilevel modeling is appropriate. There was significant variation in teacher-level, school-level and country-level mean frequencies of use of inquiry-based methods of instruction ($\sigma^2 = 0.190$, $p<0.01$; $\tau_\pi = 0.291$, $p<0.01$; $\tau_\beta = 0.003$, $p<0.01$). The analysis of intraclass correlations indicated that 39% of variation in the dependent variable is attributed to differences at the individual level, 60% of variation – to differences at the school level, and only 1% of variation was due to variation between countries. Such break down of variance implies that there is not much variation among countries in the use of inquiry-based methods and that modeling the third level is not necessary. However, since the study is exploratory in nature and the variation at the country level is significant, though small in magnitude, we decided to proceed with three level modeling to determine whether this variation is conditioned by the hypothesized predictor variables or it is purely random in nature. In addition to that, modeling the third level allowed toimprove estimation at Level I and Level II.

Prior to implementation of the multilevel analysis, some model assumptions have been tested. Based on the description of sampling procedures for TIMSS 2007, countries used either complex random sampling mechanisms or included the complete populations of schools in the international testing; so the assumption of random sampling of clusters was met. Figure 2 shows the distribution of Level 1 residuals. Although the distribution is not quite normal (there is a spike in the center of the distribution), there does not seem to be outliers (skewedness=0.130, se=0.005). Figure 3 shows a histogram of the difference of estimated and model implied Level II intercept. The distribution approximates normal without any outliers (skeweness=0.592, se=0.034). Figure 4 shows a similar histogram for Level III. From the histogram it can be concluded that the distribution of Level III residuals also approximates normal and there are no outliers at the level (kurtosis=-0.59, se=0.374)[†]. Overall, the data satisfies the assumption of error normality at all levels and multilevel modeling is appropriate[‡].

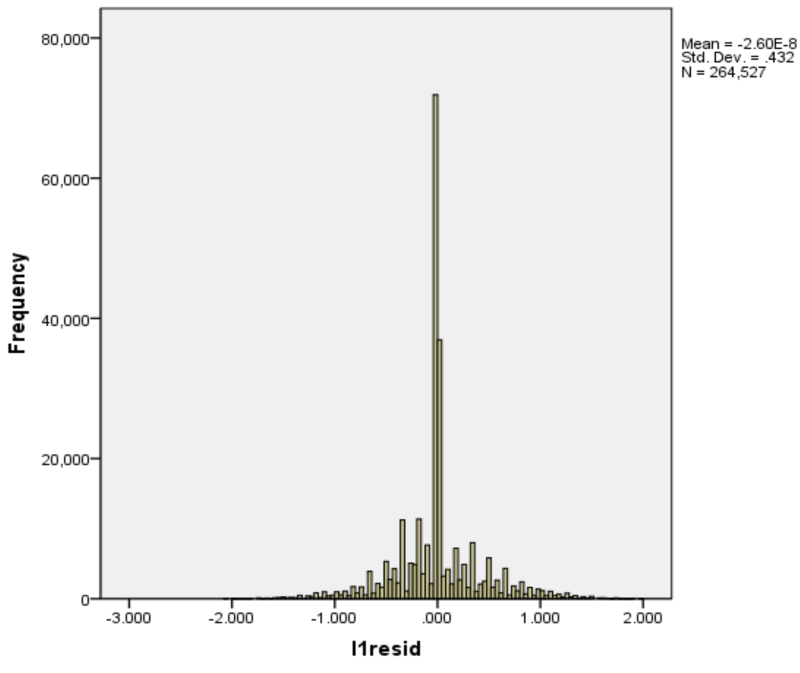

Figure 2. Distribution of Level I Residuals

[†] I could not plot chips vs. mdist because they are not available in HLM3.

[‡] I have not implemented the test of homogeneity of Level 1 error variances. This option is not available in HLM3.

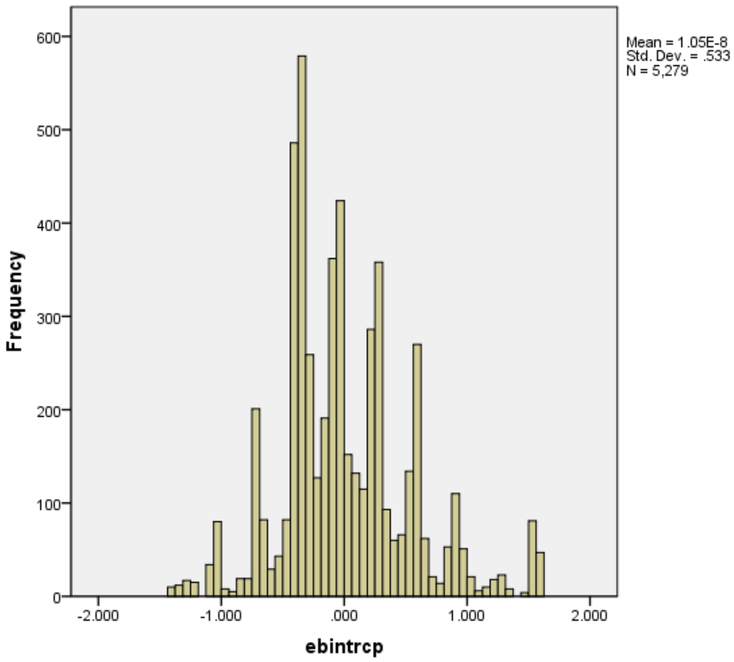

Figure 3. Distribution of Level II Residuals

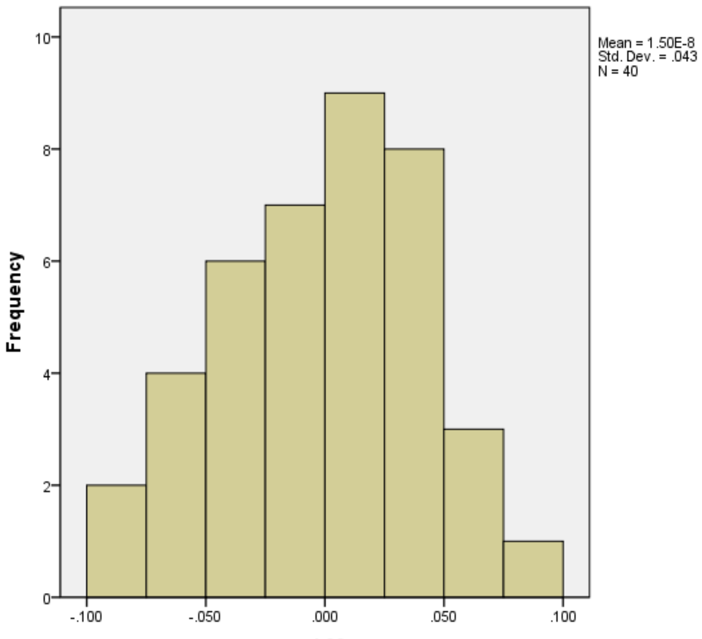

Figure 4. Distribution of Level III Residuals

The results of the deviance test indicated that the conditional model was better fitting than the fully unconditional model. The value of the test statistic was D_{uncond}-D_{cond}= 333583.685 – 333112.638 =471.047. The corresponding degrees of freedom for the test were obtained from: df_{cond}-df_{uncond} = 19 – 4 = 15. At α=0.05, the test statistic was found significant compared with χ^2critical = 25). The average conditional reliability of Level I intercepts was very high (0.979), which tells us that the model was able to successfully distinguish between teachers with the same values of all predictor variables. The average conditional reliability of Level II intercepts was also relatively high (0.516), though lower than average reliability of the intercept estimates at Level I (as should be expected). The latter reliability estimate implies that the model was also relatively successful in distinguishing between mean frequencies of use of inquiry-based techniques among schools within a country that were the same on the characteristics indicated by the predictor variables.

In testing model coefficients Bonferroni correction was utilized to adjust for Type I error. The familywise error rate was set at α=0.1. Since 34 tests were conducted, error rate per test was set at α= 0.1/34=0.003. Table 1 summarizes the results of estimation of fixed effects. The results indicate that the grand mean frequency of utilization of inquiry-based methods is around 2 (se=0.063, t(37)=39, p<0.003), which means that inquiry-based methods are used in teaching about half the lessons across the world. In addition to that, the results indicate that out of the set of potential predictors, only some of the individual level predictors have a significant influence on the extent to which a teacher utilizes inquiry-based methods (mean frequency of use of the method). Specifically, the main predictors seem to be teacher's level of experience, level of education and the number of students in the class. While the first two predictors have a small, but significant positive effect (EXPER= 0.001, t(264,511)= 9.463, p<0.003; EDU=0.008, t(264,511)=4.606, p<0.003), the latter has a small negative effect (CL_SIZE = -0.0001, t(264,511)= -17.077, p<0.003). Age might also have some positive effect on the predictor variable, with the p-value being relatively close to the level of significance (AGE=0.003, t(264,511)= 2.894, p=0.004).

Table 1. Estimates and significance test results for fixed effects

Fixed Effect	*Coefficient*	*Standard Error*	*T-ratio*	*D.f.*	*P value*
INTRCPT1, P0					
Fr INTRCPT2, B00					
INTRCPT3, G000	2.414	0.063	38.618	37	<*0.003*
GDP, G001	0.000001	0.0000001	0.588	37	0.560
STI_RANK, G002	-0.032	0.012	-2.731	37	0.010
For LOW_SES, B01					
INTRCPT3, G010	0.007	0.007	0.885	5,271	0.376
For SCH_SIZE, B02					
INTRCPT3, G020	0.009	0.011	0.766	5,271	0.444
For PDEV_CONTENT, B03					
INTRCPT3, G030	0.019	0.010	1.915	5,271	0.055
For PDEV_SKILLS, B04					
INTRCPT3, G040	-0.022	0.011	-2.070	5,271	0.038
For MAT_SHORT, B05					
INTRCPT3, G050	0.0004	0.009	0.048	5,271	0.962
For EQUIP_SHORT, B06					
INTRCPT3, G060	-0.004	0.008	-0.525	5,271	0.599
For TEACHER_SHORT, B07					
INTRCPT3, G070	-0.0002	0.008	-0.025	5,271	0.980
For AGE slope, P1					
For INTRCPT2, B10					
INTRCPT3, G100	0.003	0.001	2.894	264,511	*0.004*
For SEX slope, P2					
For INTRCPT2, B20					
INTRCPT3, G200	0.002	0.002	0.819	264,511	0.413
For EXPER slope, P3					
For INTRCPT2, B30					
INTRCPT3, G300	0.001	0.0001	9.463	264,511	<*0.003*
For EDU slope, P4					
For INTRCPT2, B40					
INTRCPT3, G400	0.008	0.002	4.606	264,511	<*0.003*
For PERCEP slope, P5					
For INTRCPT2, B50					
INTRCPT3, G500	-0.003	0.002	-2.014	264,511	0.044
For CL_SIZE slope, P6					
For INTRCPT2, B60					
INTRCPT3, G600	-0.0001	0.00001	-17.077	264,511	<0.003 -

Discussion

Two main conclusions can be made based on the results of the study. First, not all variability in the extent of utilization of inquiry-based methods can be explained by teacher-associated characteristics. Based on the results of intraclass correlation analysis, most of the variability is determined at the school level and, in addition, some slight variability is explained by differences among countries. However, the exploratory analysis of the potential predictor variables has been only marginally successful: the study was able to identify only potential predictors at the individual level. None of the explored predictors at the school and country level were significant. Conditional on the availability of international-level data, future studies should explore other potential variables at the school and country level that might explain variation in the use of inquiry-based methods. One such variable is whether school is affiliated with a religious institution and has a religious mission. Another potential variable is whether in a country church is separated from the state. Some other cultural characteristics at both school and country level could be at play.

The potential predictors identified at the individual level include teacher's age, level of experience, level of completed education, and class-size. These predictors are consistent with the findings of the prior studies by Garcia (2003) and Inda (2013), although some of the teacher characteristics, which were considered in prior studies, such as teachers' attitudes towards science and inquiry-based instruction, teachers' type of education (whether they had specialized training in science), and teachers' experiences in science during school and university education were not considered here due to availability of data in the secondary source, which was used in the study. Given the observational nature of the TIMSS 2007 study, which was used as a source of the study, no causal inferences can be made from the study. The results of the study are relatively reliable to the extent that TIMSS 2007 had a rigorous design and was administered with much attention to accuracy. Given the international nature of the data collection effort and the involvement of major research centers of Ministries of Education from around the world, one can be relatively confident that the complex scheme used for random sampling of schools within countries was implemented with adequate attention to detail.

Potential bias could have been introduced by incorrect specification of the model used for this study specifically. To test the effect of possible misspecification, some basic sensitivity analysis was conducted. It was found that Level I residuals are not correlated with any of the level I predictors, hence the assumption of independence of level I residuals from the Level I covariates was satisfied and it can be assumed that Level I model was correctly specified and produced accurate estimates of the effects. To test whether estimation of Level I coefficients was affected by misspecification at Level II the model was refitted without centering at Level I. The resulting Level I coefficients has changed very slightly (in the order of third place decimals), which, combined with the fact that Level I residuals were not correlated with Level I covariates indicates that estimates of the Level I fixed effects are relatively accurate.

No sensitivity analysis was conducted to test whether potential misspecification at Level II and Level III had any effect of estimation of the corresponding coefficients at these levels. This kind of analysis is left for future studies that should probably include the covariates at the school and country level that were suggested by this study in addition to some newly hypothesized covariates to see whether the newly specified model produces different estimates of the fixed effects and the covariates suggested by this study are, in fact, significant. Future studies might also want to consider following Raudenbush & Bryk's (2002) suggestion to include the same covariates in all intercept and slope models at the same level as a way to counteract potential effects of Level II and Level II misspecification on the estimates of the fixed effects at the levels. Another limitation of the study is that we did not deal with potential within-level and cross-level interaction effects and did not address the issue of multicollinearity. This study was largely exploratory in nature, with the expectation that more advanced analysis should be conducted once more advanced statistical software and more detailed international level data is available.

References

Alfireri, L., Brooks, P.J., Aldrich, N. J., tenenbaum, H.T. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, *3*(1), 1-18.

Azen, S. P., Van Guilder, M., & Hill, M. A. (1989). Estimation of parameters and missing values under a regression model with non-normally distributed and non-randomly incomplete data. *Statistics in Medicine*, *8*(2), 217-228.

Barthlow, M. J., & Watson, S. B. (2014). The effectiveness of process-oriented guided inquiry learning to reduce alternative conceptions in secondary chemistry. *School Science and Mathematics*, *114*(5), 246-255.

Collins, A. (1986). *A sample dialogue based on a theory of inquiry teaching* (Report No. 367). Cambridge, Mass.: Bolt, Beranek & Newman, Inc.

DeBoer, G. E. (1991). *A History of Ideas in Science Education: Implications for Practice*. Teachers college press, 1234 Amsterdam Avenue, New York, NY 10027.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society. Series B (methodological)*, 1-38.

Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching a meta-analysis. Review of Educational Research, *82*(3), 300-329.

Garcia, K. (2003). The effect of teacher attitude, experience and background knowledge on the use of inquiry-method teaching in the elementary classroom. Retrieved from https://www.utd.edu/nsm/scimathed/resources/SER/SCE5308_03/Final_PaperCG.pdf

Gautreau, B. T., & Binns, I. C. (2012). Investigating student attitudes and achievements in an environmental place-based inquiry in secondary classrooms. *International Journal of Environmental and Science Education*,*7*(2), 167-19.

Glasson, G. E. (1989). The effects of hands-on and teacher demonstration laboratory methods on science achievement in relation to reasoning ability and prior knowledge. *Journal of Research in Science Teaching*, *26*(2), 121-131.

Haitovsky, Y. (1968). Missing data in regression analysis. *Journal of the Royal Statistical Society. Series B (Methodological)*, 67-82.

Hung, M. (2009). *Achieving science, math and reading literacy for all: The role of inquiry-based science instruction* (Doctoral dissertation). The University of Utah.

Inda, A. N. (2013). Influence of teacher's characteristics on effective use of inquiry based approach in teaching science in preschools In Kuja Zone, Rongo District (Doctoral dissertation, University of Nairobi).

International Association for Evaluation of Educational Achievement. (2007). *Trends in International Mathematics and Science Study 2007*. Retrieved from http://www.iea.nl/timss2007.html

International Monetary Fund. (2010). *World Economic Outlook Database. Macroeconomic data series.* Retrieved from http://www.imf.org/external/pubs/ft/weo/2010/01/weodata/index.aspx

Kim, J. O., & Curry, J. (1977). The treatment of missing data in multivariate analysis. *Sociological Methods & Research*, *6*(2), 215-240.

Kyle Jr, W. C. (1985). What Research Says: Science Through Discovery: Students Love It. *Science and Children*, *23*(2), 39-41.

Lindberg, D. H. (1990). What goes' round comes' round doing science. *Childhood Education*, *67*(2), 79-81.

Lloyd, C. V., & Contreras, N. J. (1985, December). *The Role of Experience in Learning Science Vocabulary*. Paper presented at the Annual Meeting of the National Reading Conference, San Diego.

Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry-based science instruction—what is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of research in science teaching*, *47*(4), 474-496.

National Science Board. (1991). *Science & engineering indicators -1991*. Washington, DC: U.S. Government Printing Office. (NSB 91-1)

Narode, R., Heiman, M., Lochhead, J., and Slomianko, J. (1987). Teaching thinking skills: Science. Washington, D. C.: National Educational Association.

Pea, C. H. (2012). Inquiry-based instruction: Does school environmental context matter. *Science Educator*, *21*(1), 37-43.

Wu, P. L., Wu, L. C., Shih, W. L., & Wu, M. L. (2014). A comparison study on the influence of implementing inquiry-based instruction on science learning motivation and interest. *Asian Journal of Management Sciences & Education*, , *3*(3)` cddrteytryey.

Rakow, S. J. (1986). *Teaching Science as Inquiry. Fastback 246*. Phi Delta Kappa, Eighth and Union, Box 789, Bloomington, IN 47402.

Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Sage.

Sandoval, W. A., & Harven, A. M. (2011). Urban middle school students' perceptions of the value and difficulty of inquiry. *Journal of Science Education and Technology*, *20*(1), 95-109.

Staver, J. R. (1986, September). *The Constructivist Epistemology of Jean Piaget: Its Philosophical Roots and Relevance to Science Teaching and Learning.* Paper presented at the United States-Japan Seminar on Science Education, Honolulu.

Wagner, C. S., Brahmakulam, I., Jackson, B., Wong, A., & Yoda, T. (2001).*Science and technology collaboration: Building capability in developing countries* (No. RAND/MR-1357.0-WB). RAND CORP SANTA MONICA CA.

Wiberg, M., & Andersson, E. (2010, July). School-Effectiveness in Mathematics in Sweden compared with countries in Europe and Asia-Pacific. In *Paper submitted for presentation at the International Association for the Evaluation of Educational Achievement (IEA) Conference*.