Full Paper Proposal IRC 2010

Sampling Design for the Teacher Education and Development Study in Mathematics (TEDS-M): Challenges, Solutions, Restrictions

Sabine Meinck[1], *Jean Dumais*[2]

Abstract:

The Teacher Education and Development Study in Mathematics (TEDS-M) focuses on how teachers are prepared to teach mathematics in primary and lower secondary schools in seventeen countries. This is the first large-scale assessment in teacher education using statistical sampling. The demanding study goals, combining four target populations into one survey on the one hand, and the complexity and differences of the teacher education systems in participating countries, on the other hand, posed particular challenges to design a multi-purpose international sampling plan. The four target populations of the TEDS-M study are teacher preparation institutions, future primary and lower secondary mathematics teachers, and their educators, for which reliable estimates of their main characteristics were required. The overall achieved participation rates in most countries complied with the demanding standards set by TEDS-M and therefore ensure a high validity of the data collected. Although high technical standards were maintained, certain restrictions must be put on data analysis and on the interpretation of the results. Data observed from populations with low participation rates will be annotated as such in the forthcoming international reports. Further, specific structural or political circumstances made it necessary for some countries to implement sampling or operational procedures that deviated from the international design. Cross-country comparisons must be made with caution; results always need to be embedded into adequate contextual explanations.

Experiences gained throughout the implementation of this study can contribute valuably to the specification of sampling designs in further studies in higher education.

Keywords: *sampling, survey methodology, teacher education, TEDS-M 2008*

[1] IEA Data Processing and Research Center, Hamburg, Germany; e-mail: sabine.meinck@iea-dpc.de
[2] Methodology Branch, Statistics Canada, Ottawa, Canada; e-mail: Jean.Dumais@statcan.gc.ca

Full Proposal:

Introduction

The Teacher Education and Development Study in Mathematics (TEDS-M) is the first study of post-secondary education conducted by the International Association for the Evaluation of Educational Assessment (IEA) (Tatto *et al.*, 2008). The IEA requires that their cross-national comparative studies rely on nationally representative samples. The demanding study goals, combining four target populations into one survey, and the complexity and differences of the teacher education systems in the seventeen participating countries[3], posed particular challenges to design a multi-purpose international sampling plan.

Target Population Definitions

The study targeted (1) the institutions where future primary and lower secondary teachers are prepared to teach mathematics; (2) the educators of future teachers; (3) the future primary and (4) future lower secondary teachers themselves in their last year of training.

The definition of clearly outlined survey targets that work in all participating countries and lead to internationally comparable populations turned out to be a challenging key component of this study.

The international target population of **teacher preparation institutions** was defined as the set of secondary or post-secondary schools, colleges or universities which offer structured "opportunities to learn" (i.e., a program or programs) on a regular and frequent basis to future teachers within a route of teacher preparation (see TEDS-M Conceptual Framework [Tatto *et al.*, 2008] for key definitions). For an institution to be part of the target population, it was not necessary that it includes teaching the content of mathematics, but it was necessary that the institution includes teaching the pedagogy of mathematics (IEA, 2007).

The target population of **educators** comprised all persons with regular, repeated responsibilities to teach mathematics, mathematics pedagogy and/or general pedagogy to

[3] Botswana, Canada (4 provinces), Chile, Chinese Taipei, Georgia, Germany, Malaysia, Norway, Oman (secondary education only), Philippines, Poland, Russian Federation, Singapore, Spain (primary education only), Switzerland (German speaking parts), Thailand, United States of America (public institutions)

future teachers in one of the compulsory courses of their teacher preparation program at any year of the program active during the collection year. Considering the definitions of in-scope institutions and educators, two limitations of the TEDS-M educator survey become clear. (1), educators of mathematics belonging to institutions that offer the first but not the second phase of a consecutive route are not targeted by the survey. (2), the in-/exclusion of educators in the survey depends on the proportion of compulsory and elective courses in teacher preparation programs in a country. In Chinese Taipei for example, contrary to other participating countries, only a few of the courses for future teachers are compulsory. Instead, students have to choose most of their courses from specific packages of topic-related courses in order to receive full credits on completing their program. Therefore the nationally defined educator population in Chinese Taipei was – in relation to the sizes of the future teacher populations – much smaller than in other countries.

It should also be noted that, for various – mainly organizational – reasons, educators within an institution could not be and were not directly linked to the future teachers of that institution.

Finally, the target population of **future teachers** comprises all members of a route in their last year of training enrolled in an institution, explicitly intended to prepare individuals qualified to teach mathematics in any of the grades 1 to 8. TEDS-M distinguishes between two different groups of future teachers: future teachers who will be certified to teach to primary students and future teachers who will be certified to teach to lower secondary students. These two groups are referred to as two distinct “levels”. In some countries, the distinction between primary and lower secondary levels is not feasible within a program. For example, teachers may be prepared – and certified – for both levels because they will be expected to teach mathematics at any level from grade 1 to grade 8. Among the participating countries, Chile, Germany, Norway, Poland and Thailand offer such programs. In institutions with such programs, half of the sampled future teachers were randomly selected to answer for primary while others were to answer for lower secondary education[4].

Sampling Relevant Structural Features of Teacher Preparation

Two concepts are playing a key role in the organization of teacher preparation: the program and the route. A program is a specific pathway that exists within an institution, leading to the award of a common credential on completion. A route is a set of teacher education programs available in a given country. TEDS-M distinguished between concurrent and consecutive routes. **Concurrent** routes consist of programs that include studies in the subjects future teachers will be teaching (academic studies), studies of pedagogy and education (professional studies) and practical experience in the classroom (Tatto *et al.,* 2008). **Consecutive** routes consist of a first phase for academic studies (leading to a degree or diploma), followed by a second phase of professional studies and practical experience (leading to a separate credential / qualification); the first and second phases need not be completed in the same institution (Tatto *et al.*, 2008).

Programs within a given route share a number of common features that distinguish them from teacher preparation programs in other routes. For example, a set of institutions in a country may offer a 5-year *concurrent* program leading to a diploma in primary education. Another (or overlapping) set of institutions in that country may offer a c*onsecutive* program consisting of a first phase for academic studies (leading to a degree, for example a bachelor of mathematics), followed by a second phase of professional studies and practical experience, leading to a separate credential that qualifies for the teacher profession.

The various unique programs offered by the institutions may share a number of features: duration, respective scopes of subject matter and pedagogy, etc.. In some countries programs with similar features are grouped under a **program type**. For example, four concurrent program types are offered in Georgia: (1) Bachelor in Pedagogy (4 years); (2) Bachelor in Pedagogy (5 years); (3) Bachelor in Mathematics; (4) Master in Mathematics. All the four program types qualify their graduates for the teacher

[4] Achievement tests of differing levels of difficulty were created for the two different future teacher target populations.

profession. While nine institutions offer a four-year 'Bachelor in Pedagogy' program, only five institutions offer the 'Bachelor in Mathematics', etc. Conversely, in some countries, where the defining features of programs are less distinctive, there may be no need to define program types.

All programs that prepare future teachers to teach mathematics to students at primary and lower secondary schools were of interest to TEDS-M.

Participating countries identified various program types with greatly varying features. While in some countries only a single program type exists that prepares, e.g., primary teachers (Chile, Chinese Taipei, Philippines, and Russian Federation); Poland specified no less than 12 different program types. Some program types prepare specialist teachers (teaching only mathematics or mathematics plus another subject) whereas others prepare generalist teachers (teaching three or more subjects). Moreover, different program types prepare future teachers to teach to a great variety of grade spans. For example, while most primary teachers in Germany are prepared to teach to grade 1 to 4 only, future teachers in Norway and Chile are certified to teach mathematics to students up to grade 8 or even grade 10. Further, in many countries programs qualify future teachers to teach to lower secondary and also to upper secondary students. This means, that programs preparing teachers for the grades below 9 as lower grade span were in scope of the study, while programs preparing teachers for upper secondary only (grade 9 and above) were not in scope of TEDS-M, and would consequently not contribute to estimates describing features of the population of future lower secondary teachers. The variety of program types specified by participating countries will be detailed in the technical report of this study (Meinck & Dumais, 2010).

Consequently, simple cross-country comparisons across different program types seem to be ambiguous and have to be conducted with great care, embedding the results into well-explained contexts.

It should be noted that the structural features of teacher preparation within participating countries were mostly unknown when defining the national sampling plans, because official statistical information on these features was rarely available. In fact, in many countries, the programs offered in the various institutions became known only during the implementation of the survey. Last-minute adaptations to the sampling plans, developed over extensive discussions between the national teams, the study center and the sampling team, and a meticulous documentation of any changes were indispensable to maintain the validity and high quality of the national samples.

In forthcoming studies with similar target populations, it might be wise to insist that the National Research Coordinators explore thoroughly the features of their education systems and assess the applicability of the intended definitions to their national contexts before the actual survey is conducted. Though adding costs at the beginning of a survey, this may reduce costs later on and contribute to the success of the study.

International Sampling Design and National Implementations

In short, the international sampling plan implemented in TEDS-M is a stratified multi-stage probability sampling design. Random samples were required for each population. The targeted individuals (the future teachers and the educators) were randomly selected from a list of in-scope educators and future teachers for each of the randomly selected teacher preparation institutions. Further details are given in Dumais & Meinck (2010a).

The minimum sample sizes set were aiming for an effective sample size of 400 future teachers per level and route. In order to achieve this goal, a minimum of 50 institutions per level and route was to be selected. The actual number of future teachers required for each level and route within the selected TP institutions and overall was dictated mainly by the total number of institutions in the country, the size of the institutions in the country, and the selection method used within the institutions.

Teacher preparation institutions that offer education both to future primary and to lower secondary school teachers could be part of both samples. Similarly, institutions that offer more than one route to students could be part of more than one sample.

Among the seventeen countries participating in TEDS-M, twelve identified fewer than 50 (or only slightly more than 50) eligible institutions. These countries conducted a census of institutions[5]. Therefore, in these countries, the sample design can no longer be described as a two-stage cluster design; the design became a stratified simple random

[5] Although Chinese Taipei belonged to countries with less than 50 institutions, a sample of two institutions per level was selected from a number of very small institutions.

sample. The effect of the change in designs can be seen in the high precision of the estimates for such samples.

If an institution offered different programs to become a primary/lower secondary mathematics teacher, then all the programs were automatically selected and the eligible future teachers inside that institution were stratified by these programs. Within the programs, a sample of future teachers was selected (minimum sample size of 30 future teachers), either through individual future teacher sampling or the selection of full session groups. Session groups were, in this context, defined as mutual exclusive and exhaustive groups of future teachers within a program. The latter within-institution sampling method was widely used only in three countries (Chinese Taipei, Germany, and Russian Federation). Because of the expected increased clustering effect, the intended sample sizes were set higher.

An 'institutional program questionnaire' was to be completed for each identified program eligible to TEDS-M in a selected institution. Further, a minimum of 30 mathematics and 30 mathematics pedagogy educators was selected within each participating institution. This requirement led to a census of educators in 95% of all institutions participating in the study.

The within-institution sampling was in the responsibility of the national study centers. Specific software provided by the IEA Data Processing and Research Center, called "WinW3S" was to be used to perform the sampling of future teachers and educators. This software stored the selection probabilities and the participation status of each individual in a database that were later on used to calculate design and estimation weights for all survey units.

Because of the broad variety of teacher education systems in the participating countries in terms of structure and size, each national sampling plan is unique, ranging from a stratified multi-stage probability sampling plan with unequal probabilities of selection (Chinese Taipei, Philippines, Russian Federation, Spain and the United States of America) to a simple and complete census of all units of interest (Botswana, Chile, Georgia, Norway, Oman, Singapore, Thailand). The remaining countries (Canada, Germany, Malaysia, Poland and Switzerland) conducted a census of institutions but selected samples within institutions. Each national sampling plan will be detailed in the upcoming technical documentation of this study.

Overall, more than 500 institutions, 15,000 future primary teachers, 8,500 future secondary teachers, and 5,000 educators participated in this study.

The various sampling designs, implemented in dramatically different settings and conditions, varied significantly in their achieved precision. The efficiency of sampling designs can be measured by the *design effect*. This discussion will be subject of a forthcoming paper.

Modifications of the International Sampling Plan

One important modification of the international sampling plan was the reduction of scope of the national implementation. Countries could choose to reduce their target populations for political, organizational or operational reasons that would otherwise have made it extremely difficult to obtain complete national coverage. For example, the United States covered only public institutions that account for about 60% of the total population of future teachers, but surveyed future teachers in private institutions one year later as subject of a separate national survey (Babcock *et al.*, 2010). Reduced coverage means that the survey results cannot be deemed representative of the entire national teacher education system in target of TEDS-M. Exclusions or reduced coverage could not amount to more than 5% of the internationally defined target populations. Higher rates will be annotated in the upcoming international reports. Table 1 gives an overview about coverage and exclusions in participating countries.

Two further particularities of national sampling designs require attention.

In Germany, the specific two-phase structure of teacher education made it necessary to define the term ‘teacher preparation institutions’ (and therefore the respective primary sampling unit) for the future teachers and educators populations in different ways. It needs to be understood that due to the structure of the sample (and the German system of teacher education), data collected for the populations of educators and future teachers are not connected in any way.

Table 1: Exclusions and Coverage

Country	Exclusions	Coverage
Botswana	None	100% in all target populations
Canada (4 provinces)	None	Participating provinces: Ontario, Quebec, Nova Scotia, Newfoundland & Labrador; within the 4 provinces 100% in all target populations
Chile	Future teachers on practicum in remote areas (2% of institutions; 2% of educators; 3.8% of future primary teachers; 3.6% of future lower secondary teachers)	100% in all target populations
Chinese Taipei	Very small institutions (26.1% of institutions; <4% of educators; 4.5% of future primary teachers; 4.7% of future lower secondary teachers)	100% in all target populations
Georgia	Russian and Azeri sectors of institutions (1.4% of future primary teachers; 1.7% of future lower secondary teachers)	100% in all target populations
Germany	One federal state, very small session groups (6% of institutions offering primary education and 3.7% of future primary teachers; 7% of institutions offering lower secondary education and 5.6% of future lower secondary teachers) Very small institutions (22% of institutions eligible for the educator survey; <5% of educators)	100% in all target populations
Malaysia	None	Due to low participation, program type 'Bachelor of Education in Teaching of English as Second Language with minor in mathematics' not covered.(<5% of future primary teachers)
Norway	None	100% in all target populations
Oman	None	100% in all target populations (did not participate in the study of future primary teachers because such future teachers were not educated during data collection period)
Philippines	Very small institutions (7.4% of institutions; <5% of educators; 2.1% of future primary teachers; 1.7% of future lower secondary teachers)	100% in all target populations
Poland	Very small institutions (3.8% of institutions; <5% of educators; 3.0% of future primary teachers; 0.4% of future lower secondary teachers)	Institutions with consecutives programs only were not covered (8.5% of institutions; percentage of not covered educators unknown; 23.6% of future primary teachers; 29.0% of future lower secondary teachers)
Russian Federation	None	Secondary Pedagogical Institutions (amount unknown)
Singapore	None	100% in all target populations
Spain	None	Only institutions offering education to future primary teachers covered (populations of institutions, future primary teachers, and educators are affected)
Switzerland	None	German speaking parts covered only (all populations are affected)
Thailand	None	100% in all target populations

USA	None	public institutions covered only

Norway offers four different program types of primary/lower secondary teacher preparation. The clienteles of program types ‘ALU - general teachers for primary and lower secondary’, ‘ALU - general teachers for primary and lower secondary with special program for mathematics’, and ‘Master - teachers in lower and higher secondary school’ are partly overlapping. Therefore analysis across these program types is inappropriate. It is strongly recommended to conduct analysis only separately by each program type. Further it should be noted that future teachers from one program type (ALU - general teachers for primary and lower secondary) could not be reached in their final year because a major percentage of them spend this time outside the institutions. Therefore, future teachers from this program were tested in deviation to the international study design at the time when they were taking their compulsory mathematics courses. This could be within their 4th or 6th semester.

Participation Rates

The overall achieved participation rates in most countries complied with the demanding standards set by TEDS-M and therefore ensure the validity of the data collected. In particular, 11 out of 16 countries participating in the survey of future primary teachers achieved participation rates above 75% (i.e., combined participation rates for institutions and future teachers within participating institutions). The same can be said for the future secondary teacher population for 10 out of 16 participating countries. Looking retrospectively, the development of strategies to ensure high participation rates, implementing experiences from the field test[6] was a critical factor influencing the success of the study in most participating countries. In fact, out of ten countries participating in the field test (and monitoring response rates) nine were able to increase the response rates of future primary teachers comparing to main survey results, even though they had selected a very small convenience sample for the field test. Table 2 displays the participation rates achieved for the four different target populations.

[6] A field test is conducted in all IEA studies about one year before the main data collection. Its main purpose is the test of the items and all survey operation procedures on a small sample.

Data observed from populations with low participation rates (below 75% combined) will be annotated in the forthcoming international reports.

Table 2: Participation Rates (%)

Country	Institutions (completion of IPQs)*	Combined Participation Rates**		
		Future Primary Teachers	Future Lower Secondary Teachers	Educators
Botswana	100	86	88	98
Canada (4 provinces)	37	5	21	26
Chile	88	68	63	54
Chinese Taipei	100	90	97	95
Georgia	100	77	67	97
Germany	100	76	81	56
Malaysia	57	93	72	57
Norway	96	63	58	Could not be calculated
Oman	100	n.a.	93	85
Philippines	85	75	83	80
Poland	86	68	69	68
Russian Federation	91	91	92	91
Singapore	100	90	91	85
Spain (primary education only)	96	78	n.a.	85
Switzerland (German speaking parts)	94	76	81	52
Thailand	96	97	96	88
USA (public institutions)	83	71	69	14

*An institution counted as participating for this part of the survey if it submitted at least one institutional program questionnaire
**Participation rate of institutions multiplied by response rate within participating institutions

Weighting and Variance Estimation

Given the complexity of the survey design, estimation weights were calculated to reflect the different selection probabilities at the various sampling stages and non-response adjustments (for more details, see Dumais & Meinck, 2010b). For the correct estimation of sampling errors, the TEDS-M data base will provide replicated weights to apply Fay's variant of Balanced Repeated Replication (BRR) (McCarthy, 1966; Fay, 1989; Judkins, 1990). Only the usage of the correct weights and estimation method allows the production of unbiased country-level estimates from the observed sample data (Lohr, 1999). Statistical software packages that feature BRR (e.g., the IDB Analyzer[7] or WesVar[8]) have to be used to obtain correct estimates for sampling variances.

Conclusions

The variety of teacher education in the countries participating and specific circumstances leading to inevitable variations or adaptations from the international sampling plan entail being cautious when analyzing the data, interpreting the results or conducting cross-country comparisons. Results have to be embedded into adequate contextual explanations, leading to carefully drawn conclusions. In fact, the complexity of the survey makes secondary analysis more difficult than for other large scale assessment datasets, like TIMSS[9], for example.

The public use international database of the TEDS-M study will however deliver a precious basis to retrieve evidence about effective teacher preparation practices and mechanisms under differing cultural contexts, when carefully regarding the specific structural features of the collected data as well as limitations of analyses results.

Experiences gained throughout the implementation of this study can contribute valuably to the specification of sampling designs in further studies in higher education.

References:

Babcock, B., Babcock, P., Buhler, J., Cady, J., Cogan, L., Houang, R., Kher, N., Rosolova, K., Schmidt, W.H., Wight, K. (2010). Breaking The Cycle: An International Comparison of U.S. Mathematics Teacher Preparation, Michigan State University Center for Research in Mathematics and Science Education, http://usteds.msu.edu.

Dumais, J., Meinck, S. (2010a). Sampling Design. - In: Tatto, M. T., Schwille, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (2010): *Teacher Education and Development Study in Mathematics (TEDS-M): Technical Report*. East Lansing,

[7] Available at http://www.iea.nl/iea_studies_datasets.html (free of charge)
[8] Available at http://www.westat.com/westat/statistical_software/wesvar/wesvar_downloads.cfm
[9] Trends in International Mathematics and Science Study, conducted by IEA

MI: Teacher Education and Development International Study Center, College of Education, Michigan State University (in preparation).

Dumais, J., Meinck, S. (2010b). Estimation Weights, Participation Rates, and Sampling Error. - In: Tatto, M. T., Schwille, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (2010): *Teacher Education and Development Study in Mathematics (TEDS-M): Technical Report*. East Lansing, MI: Teacher Education and Development International Study Center, College of Education, Michigan State University (in preparation).

Fay, R.E. (1989). Theoretical Application of Weighting for Variance Calculation, *Proceedings of the Section on Survey Research Methods of the ASA*, 212-217.

IEA (2007). TEDS-M Sample Preparation Manual, Hamburg: IEA-DPC.

Judkins, D.R. (1990). Fay's Method for Variance Estimation, *Journal of Official Statistics*, Vol.6, No. 3, 223-239.

Lohr, S. (1999). *Sampling: Design and Analysis,* Duxbury Press, New York.

McCarthy, P.J. (1966). Replication: An Approach to the Analysis of Data from Complex Surveys. *Vital and Health Statistics,* Series 2, No. 14, Hyattsville, MD: National Center for Health Statistics.

Meinck, S., Dumais, J. (2010). Characteristics of National Samples: Implementations of the International Sampling Design in Participating Countries. - In: Tatto, M. T., Schwille, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (2010): *Teacher Education and Development Study in Mathematics (TEDS-M): Technical Report*. East Lansing, MI: Teacher Education and Development International Study Center, College of Education, Michigan State University (in preparation).

Tatto, M. T., Schwille, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (2008). *Teacher Education and Development Study in Mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics. Conceptual framework.* East Lansing, MI: Teacher Education and Development International Study Center, College of Education, Michigan State University.