Abstract Title Page

Title: The differential effectiveness of the M@t.abel teacher professional development program in mathematics in Italy: Evidence from a random assignment evaluation

Authors:

Argentin, Gianluca, Università di Milano Bicocca, Italy, g.argentin@campus.unimib.it
Cipollone, Piero, World Bank, pcipollone@worldbank.org
Martini, Alberto, Università del Piemonte Orientale, amartini@prova.org
Pennisi, Aline, Ministero dell’Economia e delle Finanze, aline.pennisi@tesoro.it
Vidoni, Daniele, INVALSI, daniele.vidoni@invalsi.it
Abstract Body

Background / Context:

Research has proven that teachers have a fundamental influence on student results (Scheerens and Bosker, 1997; Scheerens, 2000; Scheerens, 2007). Moreover, effective teacher professional development is one of the key mechanisms for improving student achievement (Darling-Hammond, Holtzman, Gatlin & Heilig, 2005; Darling-Hammond, 2000; Kennedy, 1998). By the most recent OECD definition (TALIS survey, OECD, 2009, p. 49), “Professional development is defined as activities that develop an individual’s skills, knowledge, expertise and other characteristics as a teacher”. The M@t.abel program suits this definition both in terms of increasing teacher subject knowledge and in terms of providing math teachers with alternative solutions and methods for presenting usual contents. Indeed, M@t.abel is focused on teaching students to solve real life problems using mathematical tools and concepts, rather than learning abstract formulas and ideas. It is addressed to math teachers in grades 6-8 (middle school) and 9-10 (first two years in high school). The program is based on formal and on-line tutoring, on the application of specific teaching materials in the classroom, and on a virtual community of teachers exchanging views through on-line forums and discussion groups.

Given that the effects of any professional development program on student achievement are mediated by teacher actual practices in the classroom (Cohen & Hill, 2000; Cohen, Raudenbush & Ball, 2002), it is necessary to verify whether teachers enrolled in the M@t.abel program do follow the training, whether they apply the acquired knowledge in the classroom, and whether they change their attitudes and teaching methods. The study will address these issues by collecting information through specific questionnaires and the teacher’s M@t.abel log-book.

The difficulty in measuring the effectiveness of teaching is that “teachers have powerful effects on reading and mathematics achievement, though little of the variance in teacher quality is explained by observable variables, such as education or experience” (Rivkin, Hanushek and Kain, 2005). The study rules out issues resulting from hidden heterogeneity by designing ex ante a randomized experiment, which allows us to compare the results of two statistically equivalent groups of participants in the most objective way possible.

Purpose / Objective / Research Question / Focus of Study:

In many countries the drive to improve education has triggered a season of rigorous research on what kind of instructional practices, curricula or specific interventions do work. Italy is lagging behind in this effort due to limited availability of data on student outcomes, historic aversion to standardized testing, and lack of a tradition of evidence-based policy evaluation. The weaknesses shown by Italian students in international tests on mathematics and science (IEA-TIMSS and OECD-PISA studies), has produced a flurry of initiatives to help schools and teachers improve student achievements. Among these initiatives, teacher professional development programs supported by the Ministry of Education are playing an important role. One of these programs, called M@t.abel, is aimed at training teachers in the application of mathematics to everyday life situations. M@t.abel classroom materials have recently been expanded to cover a substantial fraction of the middle-school curriculum and the program is
being strongly promoted in four regions of Southern Italy, thanks to European Union funding. These regions are characterized by the low levels of math achievement. The key evaluation question is whether exposure to M@t.abel has a discernible effect on student achievement, both in terms of the whole population and with respect to specific subgroups of students identified in terms of prior ability and individual characteristics. The study attempts to answer this question through random assignment (at school and classroom level) and involves the collection of a large quantity and variety of data on teachers, students, and schools. In addition to the students’ math scores in standardized tests, we measure students’ attitudes towards math.

Corollary to the main question, we also investigate whether exposure to the program leads to measurable changes in teachers’ attitudes towards teaching math and in their reported teaching behaviours.

Being (to our knowledge) the first random assignment evaluation attempted in the Italian school system, the study also contributes to understanding how rigorous evaluation techniques can be carried out in Italy, what the reactions of the educational community are to practices such as random assignment, and how policy makers can be involved.

Setting:

Italy is administratively organized in 20 regions, 4 of which (Campania, Calabria, Sicilia, Puglia) receive European Union Structural Funds. These regions are characterized by poor socioeconomic characteristics, relatively weak infrastructures, low student achievement, and high rates of student drop-out. Although the M@t.abel training is available nationwide, EU funds allow for implementing one intensive version of the program specifically in these 4 Regions. This intensive version is the object of the evaluation.

Population / Participants / Subjects:

Teacher enrolment in the M@t.abel program is on an annual basis. The reference population included in the experiment consisted of 877 teachers coming from 263 schools enrolled in 2009. The actual population for the experiment consists of 666 teachers in 175 schools identified on the basis of the following selection criteria:

- not less than 3 teachers from the same school (M@t.abel is a methodology aimed at creating a “waterfall effect;” thus, if teachers do not have the chance of cooperating, the methodology should not sort effects);
- teachers not previously exposed to M@t.abel training;
- a minimum number on enrolled teacher in every province to ensure a sufficient number of course participants.

After dropouts, the immediate-treatment group comprises roughly 100 schools, for a total of 400 classes and 8,000 students. The control-by-waiting group comprises about 50 schools, for a total

---

* These funds are targeted to supplementary actions aimed at improving the socioeconomic situation of the region and are available to European Regions whose annual mean per capita income is lower than 75% of EU annual mean per capita income.
of 184 classes and roughly 4,000 students. The design should allow us to detect a minimum detectable effect size of 0.20, with a power of 0.8. Within the immediate-treatment group, roughly 50% of teachers were “compliers,” while the rest were “non compliers” (i.e. these teachers enrolled in the training, but for some reason did not complete all the requirements to obtain the certification).

84% of teachers included in the study are females with an average age of 52. The youngest teacher included is 23, while the oldest is 67. Most of them (63.3%) are Biology graduates, 14.2% hold a BA in mathematics and 9.9% hold a BA in natural sciences. On average, teachers have 25 years work experience and 23 years of experience teaching mathematics in lower secondary schools. 98% has a permanent contract. As shown in table 1 no differences exist among treated and control teachers. Both groups are characterized by optimistic consideration on the possibilities of teaching and learning mathematics, even if the generally low level of student competences is a concern. Traditional representations of mathematics learning (innate male superiority, need of learning by heart rules and formulae) are challenged by both groups. Moreover, the groups are equivalent both in terms of average scoring in the various items, but also in terms of disagreement with the mean position of the group: indeed, the standard deviation is always the same both in treated and control groups. This is a further proof of the equivalence of the two groups, not only in terms of socio-economic and individual characteristics, but also in terms of declared behavior and attitudes towards teaching mathematics.

-- please enter Table 1. About here --

**Intervention / Program / Practice:**

M@t.abel is a teacher training program offered to Mathematics teachers of grade 6 to 10. In terms of the current research, we focus on training provided to teachers of grade 6-8 as these grades correspond to the middle school cycle.

Training is offered for one school year and is not repeated. The program administration in a blended, as it includes both meetings in person (8 along the year, 50h work) and online forums.

In the program, teachers receive training on specific didactic units that allow them to use alternative methods for teaching curricular contents. Each unit involves the use of practical examples and laboratories. Of the over 200 available didactic units, teachers must experiment in class 4 units.

Individual and collective reflection of the class experiences is solicited and favored by means of online forums and the discussion with the virtual class. Moreover, as a requirement for obtaining the certification, teachers must fill out an “on board diary” with details on the activities done in class, problems encountered, and student reactions to the materials.

**Research Design:**

The paper presents the design and the results after the first year of activity of an ongoing 3-years study on the effectiveness of the M@t.abel program. The basic evaluation design sees
schools and teachers within schools enrolling on a voluntary basis for the 2009/10 school year. The schools are randomized into two groups: one group receives the specialized training beginning in year 2009-10, the other group is delayed admission for one year, then released into treatment. Randomization and the delayed treatment allow a credible estimate of the effect of the first year of involvement in M@t.abel, as shown in the Figure 1.

Effects on student will be measured over three years by using the standardized tests produced by INVALSI, the Italian National Institute for Educational Evaluation. Figure 2 shows the second year effect estimation procedure. Only the results on the first year effects shall be presented at this stage.

-- Please enter Figure 1 about here --

-- Please enter Figure 2 about here --

Data Collection and Analysis:

With the support of INVALSI, the project team is collecting a large amount of primary data on both students and their teachers:

- **measure of competences and skills in mathematics**: standardized tests are being produced to assess math knowledge and skills in 7th and 8th grade. The tests for 6th grade are those already adopted by INVALSI for the yearly national assessment of student achievement. The tests for the two additional grades are anchored to those of the 6th grade and all measured on a common scale by using Rash analysis;
- **background student information**: a general background questionnaire on the social, economic and previous school experience of the students was submitted to each student alongside the standardized tests;
- **students’ attitude towards mathematics**, collected through some items in a dedicated questionnaire;
- **administrative information**: end-of-the-school year data on student grades, absences during the year and information on the school context (including whether other programs aimed at improving math achievement are carried out at the same time or have been carried out in the past);
- **teachers’ background information**: CATI is used to submit to the participating teachers a questionnaire aimed at collecting information on their demographic characteristics, education and professional career, expectations and degree of involvement in school activities other than teaching;
- **teachers’ attitudes and reported teaching behaviours pre and post treatment**: information collected through longitudinal CATI, in January 2010 and October 2010;
- **degree of teacher involvement in the program and use of the M@t.abel experience**: a pre/post teacher questionnaire was used to collect data before and after the program on teaching methods, grading methods, attitudes towards teaching mathematics and appreciation of the contents and approach used by the M@t.abel program;
• **qualitative information on the program**: a grid for in-depth interviews to a selected group of teachers and M@t.abel tutors is to be carried out at the end of the training program;
• **information about the school managers**: age, sex, years of experience, years spent in the specific school and previous schools’ enrolment in a large number of projects aimed to the development of student maths skills.

Provided the hierarchical structure of the data, HLM techniques will be used for the analyses. Moreover we will investigate the M@t.abel effects using quantile regression models (for a recent review of some typical applications of quantile regression, the reader is referred to Yu et al. (2003).

Using quantile regression models are their flexibility to allow the covariates to have different impacts at different points of the distribution and the robustness to departures from normality and skewed tails (Mata and Machado, 1996). These latter features are quite typical when looking at educational variables and at the effect that any specific teacher training program could have on any given student population.

The data shall be analyzed following these steps:
• Using student results to estimate student ability by means of IRT modeling;
• Investigating the effect and comparing it among different grades (6, 7, 8);
• Developing comparison of the effect size among students’ subgroups (gender, social background, previous maths school mark);
• Estimating the potential effect on students’ attitude toward maths;
• Estimating the potential effect on teachers’ beliefs and reported teaching behaviors.

**Findings / Results:**

Provided that the standardized student achievement tests were administered in late May and that we are still polishing the database, at the moment it is not possible to draw conclusion on the effectiveness of the program on students.

The results at student level shall be available by the end of the year.

**Conclusions:**

At this stage, we cannot yet derive any conclusion in terms of the effectiveness of the program. Yet, the simple planning of the evaluation has already achieved some effects, which - up to now - are twofold:
• Developing ex ante the program evaluation design has contributed to a substantial redesign of the program design itself (i.e. reinforcing the actual treatment in quantity and quality, encouraging school-level rather than teacher-level participation, synchronization of the professional development program with the school year)
• Initiating a large scale randomized experiment within the Italian school system has raised some interest in evidence-based research in the school context.
Appendices.

Appendix A. References


OECD (2009), Creating Effective Teaching and Learning Environments - First Results from TALIS, Paris: OECD


Appendix B. Tables and Figures

1. Mean agreement of teachers interviewed with statements on mathematics and mathematics teaching. Learning by group in M@t.abel population (min= 1, max= 10, mean, standard deviation)

<table>
<thead>
<tr>
<th>Statement</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treated</td>
<td>Controls</td>
</tr>
<tr>
<td>Learning mathematics could be easy for the students, it depends on</td>
<td>8,4</td>
<td>8,5</td>
</tr>
<tr>
<td>how it is taught</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mathematics could be made interesting even for students who are not</td>
<td>7,6</td>
<td>7,7</td>
</tr>
<tr>
<td>interested</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Many students have problems with logic operations that involve abstraction</td>
<td>6,7</td>
<td>6,8</td>
</tr>
<tr>
<td>Many students have a weak basic competences in Mathematics</td>
<td>6,7</td>
<td>6,8</td>
</tr>
<tr>
<td>A few months are enough to understand which students are gifted in</td>
<td>5,8</td>
<td>5,6</td>
</tr>
<tr>
<td>mathematics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Many students are able to connect mathematical problems with everyday</td>
<td>5,7</td>
<td>5,4</td>
</tr>
<tr>
<td>life issues</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Those who are not good in mathematics refuse studying it</td>
<td>5,1</td>
<td>5,2</td>
</tr>
<tr>
<td>Studying mathematics require natural predisposition to logical reasoning</td>
<td>5,1</td>
<td>5,2</td>
</tr>
<tr>
<td>Students obtaining best results are those who need to study the least</td>
<td>4,8</td>
<td>4,8</td>
</tr>
<tr>
<td>It’s very difficult that people not gifted in mathematics obtain good</td>
<td>4,4</td>
<td>4,3</td>
</tr>
<tr>
<td>results</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mathematics requires many hours of individual solitary study.</td>
<td>4,3</td>
<td>4,3</td>
</tr>
<tr>
<td>Studying mathematics requires learning by hearth many rules and formulae</td>
<td>3,7</td>
<td>3,9</td>
</tr>
<tr>
<td>Mathematics is a subject that is difficult to relate to students’ everyday</td>
<td>3,7</td>
<td>3,7</td>
</tr>
<tr>
<td>life</td>
<td></td>
<td></td>
</tr>
<tr>
<td>With respect to females, males can do mathematics more easily</td>
<td>3,2</td>
<td>3,2</td>
</tr>
</tbody>
</table>
**Figure 1. M@tabel – first year effect estimation**

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Immediate exposure to M@tabel</th>
<th>Delayed exposure to M@tabel</th>
</tr>
</thead>
<tbody>
<tr>
<td>1º</td>
<td>3º media</td>
<td>3º media</td>
</tr>
<tr>
<td>2º</td>
<td>2º media</td>
<td>2º media</td>
</tr>
<tr>
<td>3º</td>
<td>1º media</td>
<td>1º media</td>
</tr>
<tr>
<td>4º</td>
<td>1º media</td>
<td>1º media</td>
</tr>
<tr>
<td>5º</td>
<td>1º media</td>
<td>1º media</td>
</tr>
</tbody>
</table>

![Exposed to M@tabel](image1)

![Not yet exposed to M@tabel](image2)

**Figure 2. M@tabel – second year effect estimation**

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Immediately treated</th>
<th>Treatment delay</th>
<th>8th grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>1º</td>
<td>2009/10 2010/11 2011/12</td>
<td>2009/10 2010/11 2011/12</td>
<td>8th grade</td>
</tr>
<tr>
<td>2º</td>
<td>7th grade 8th grade</td>
<td>7th grade</td>
<td></td>
</tr>
<tr>
<td>3º</td>
<td>6th grade 7th grade</td>
<td>7th grade</td>
<td></td>
</tr>
</tbody>
</table>

**testing the assumption of cohort equivalence:**

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Immediately treated</th>
<th>Delayed treatment</th>
<th>8th grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>1º</td>
<td>2009/10 2010/11 2011/12</td>
<td>2009/10 2010/11 2011/12</td>
<td>8th grade</td>
</tr>
<tr>
<td>2º</td>
<td>7th grade</td>
<td>8th grade</td>
<td></td>
</tr>
<tr>
<td>3º</td>
<td>7th grade</td>
<td>7th grade</td>
<td></td>
</tr>
</tbody>
</table>