Hussain, S., Lindh, J., & Shukur, G. (2006). The effect of LEGO Training on Pupils' School Performance in Mathematics, Problem Solving Ability and Attitude: Swedish Data. *Educational Technology & Society*, 9 (3), 182-194.

The effect of LEGO Training on Pupils' School Performance in Mathematics, Problem Solving Ability and Attitude: Swedish Data

Shakir Hussain
Department of Primary Care & General Practice, University of Birmingham, UK
s.hussain@bham.ac.uk

Jörgen Lindh
Jönköping International Business School, Sweden
lijo@jibs.hj.se

Ghazi Shukur
Jönköping International Business School and Centre for Labour Market Policy (CAFO), Department of Economics and Statistics, Växjö University, Sweden
shgh@jibs.hj.se
ghazi.shukur@vxu.se

ABSTRACT

The purpose of this study is to investigate the effect of one year of regular "LEGO" training on pupils' performances in schools. The underlying pedagogical perspective is the constructivist theory, where the main idea is that knowledge is constructed in the mind of the pupil by active learning.

The investigation has been made in two steps. The first step was before the training and the second after training. For both cases we have constructed and included control groups. A logistic model is applied to the data under consideration to investigate whether the LEGO training leads to improving pupil's performance in the schools. To analyse the opinion studies, GLM for matched pair models and the Quasi symmetry methods have been used. Preliminary results show better performances in mathematics for the trained group in grade five, and pupils who are good at mathematics tend to be more engaged and seem to be more successful when working with LEGO.

The study has also shown that pupils have different learning styles in their approach to LEGO training. The role of the teacher, as a mediator of knowledge and skills, was crucial for coping with problems related to this kind of technology. The teacher must be able to support the pupils and to make them understand the LEGO Dacta material on a deeper level.

Keywords

Robotic Toys, Problem Solving, Constructivism, Multilevel Modelling

Introduction

This research project called "Programmable construction material in the teaching situation", aims at studying the pedagogical effects caused by the application of LEGO Dacta materials in some schools in the area of middle Sweden. For that purpose, we took two groups of study in order to make comparisons. One of them is called the control group (CG) formed by schools that have not been subjected to the experimental work with LEGO Dacta and the other called experimental group (EG) that was represented by schools that have participated pedagogically with the LEGO Dacta proposals and technological materials.

The starting point was a former project called INFOESCUELA, which was a pilot project that began in Peru 1996 (Iturrizaga, 2000). It was initiated by the Peruvian Ministry of Education, with the objective of introducing technology to primary schools through the use of LEGO Dacta materials. Over the course of three years, from 1996 to 1998, the project was expanded to cover 130 schools throughout the country of Peru. In general terms, the result of the Peruvian project showed empirical evidences that the pupils from the experimental group successfully fulfilled their school activities due to the use of LEGO Dacta material. In the referred case, the EG gained better percentages of achievement than the CG in the tests of mathematics, technology, Spanish and eye to hand coordination. The results of this research were so encouraging that LEGO Dacta was interested in studying the effects of using their products in Sweden, which represented a different school context than in Peru.

We were conscious of many of the dissimilarities between the school contexts of Peru and Sweden, and especially the more complex evaluation process in the Swedish school environment. In Peru the chances were good to find control groups of pupils that never have been in touch with computers. Due to an over all higher technological level in Sweden, we could not expect to find pupils that were totally inexperienced with computers, because most children come into contact with computers at early ages in their families and elsewhere in the community. The Peruvian research project was in that respect of a more simple nature, and the results could not easily be applied to other contexts, i. e. the circumstances of the Swedish case were more subtle.

Theoretical background

The theoretical foundation for our research project is both the theory of constructivism, deriving from Piaget's view of learning (Piaget & Garcia, 1991), and the theory of situated cognition (Lave, 1988). Constructivism in terms of Piaget is normally called individual constructivism, ie. that individual persons construct knowledge through interaction with their environment. Papert in collaboration with Piaget led him to consider using mathematics in order to understand how children learn and think. Papert believed that the computer was a tool that could allow children to explore mathematics and other curricular subjects that support constructionist learning, ie. an educational method which is based on the constructivist learning theory. This method stresses the role that concrete objects play in the complex process of knowledge construction (Papert, 1980, 1993). Relating this to pupils' interaction with the programmable construction material, it means that the pupils' individual knowledge develops continuously in building constructions of LEGO material. In situated cognition the importance of the situation or the context of learning is emphasized. Individuals participate and become absorbed in social and cultural activities. Learning is a part of these activities, which contributes meaningful knowledge to its carrier. Communication is a key concept since it mediates between the individual and her context (social situation).

Our work with the programmable material in the school classes was performed as group work. This put a lot of "attention to what kind of things that surrounds the individual; especially relations between individuals, groups, communities, situations, customs, language, culture and society" (Marton & Booth, 2000, p 28). The way of working with the LEGO Dacta material in the schools pointed to knowledge development which to a great extent seemed to happen between individuals and between groups. LEGO and LOGO rely on the same pedagogical principles, even though LEGO can be considered as an extension of LOGO. Papert wanted to design a suitable computer language for children, which had the power of professional programming languages, but was easy enough for non-mathematical beginners. The name LOGO was chosen for the new language to suggest the fact that it is primarily symbolic and only secondarily quantitative (Papert, 1980).

During the same period, Marvin Minsky was encouraging children in the new MIT Children's Laboratory to learn to control a small robot which moved on the floor and could draw a record of its movement with a pen. It seemed to Papert and Minsky a logical step to integrate control of this robotic turtle into LOGO, which later was replaced by a screen turtle, but the language remained true to Papert's ideals; that it should be a tool for learning concepts like planning, problem solving, and experimentation.

Mitchel Resnick, Professor at the lab and Director of the Lifelong Kindergarten research group, compared and contrasted several projects chosen by workshop participants which featured the construction of physical objects, often built from LEGO bricks, that were controlled by the computer (Resnick, 1991). For Resnick each project was rich in personal meaning, as it symbolized something for each of the project leaders, and secondly there was great diversity among these projects. Diversity of themes showed itself through the variety of projects, none of them "typical". He concludes that while both LEGO and LOGO encourage open-ended construction, their use alone does not lead to diversity. Participants were encouraged to forget constraints when choosing a project theme.

Most of the reported LEGO projects are small, in terms of sample size of individuals and period length of study. Criticism has been raised towards these kinds of studies, as being only stories rather than research. Becker (1987) emphasizes the importance of testable consequences in research about LOGO in his discussions of advantages and disadvantages of two types of research methodologies used to study the effect of LOGO in classroom settings: the treatment methodology and computer criticism. Several researchers, including Becker (1987), have stressed that the main evidence showing that LOGO can produce measurable learning when used in "discovery" classes has been obtained in situations close to individual tutoring. In normal-sized classes, the evidence clearly shows the need for direct instruction in the concepts and skills to be learned from LOGO, as well as further direct instruction to enable students to generalize what they have learned to transfer to other situations. This is in complete opposition to Papert's conception of the discovery approach to LOGO. Pea (1985) asks: "What contribution do we get from the computer to our mental functioning?" whether the computer is only an amplifier of cognition or a reorganizer of our minds.

Moore & Ray (1999) argue for more advanced formal statistical methods for sensitivity and performance analysis in computer experiments. They suggest, for example, multivariate methods in analysis of thorough investigations. Our research project aims to combine a wide range of different kinds of methods in a one year study.

Research questions and hypotheses

Questions about the pupils learning:

- How do the pupils learn the LEGO Dacta material?
- What kind of skills and abilities are supposed to be developed when pupils are working with the material?
- Are there any differences in the pupils' way and ability to handle the material related to age or gender?

Questions about the learning context/classroom environment:

- In what way shall the context/environment be organised to best stimulate the pupils' learning of the LEGO material?
- How to integrate the material so it will be a natural part of the pupils' everyday work?

Question about the role of teacher*:*

- What kind of role is feasible for the teacher/the pedagogy to support the pupils' learning with LEGO material?

Hypotheses

For the trained pupil with our specific program the study aimed to quantify the contribution of robotic toys (LEGO) training on mathematics, problem solving ability and student attitude (MPA) toward different issues. Formulating a proper model, interpreting the parameters and quantitative assessment will be attempted.

In the project, two hypotheses were set up:

- By using LEGO construction kits, sensors and programming tools, pupils will develop better knowledge in mathematics than pupils that do not work with the material. (1)

- By using LEGO construction kits, sensors and programming tools, pupils will develop better problem solving ability than pupils that do not work with the material. (2)

Furthermore, we were also interested to measure attitudes towards different issues related to usage of LEGO, or general techniques in school (see Appendix A).

Potentially there are two main approaches to estimating the training impact of MPA. The first approach would be to design a study of trained pupils and compare them with controlled pupils that under similar circumstances did not get any training. The second approach would be to obtain good information on the pupils' change in performance based on the MPA. The performance parameters will be obtained for individual pupils before and after training. Both of these methods were used in this study, but with emphasis on the second one.

The outline of the project

In this section we explain the design and the outline of the LEGO project.

Participating schools and classes

To get a fair comparison we wanted both groups, EG and CG, to be equivalent regarding educational, social and demographic characteristics (Lindh et al., 2003). Different educational tests were applied to these groups. When recruiting “experimental” classes (EG) our ambition was to reach a balanced distribution with schools in varying geographical regions in the middle of Sweden and representation of schools in small, medium sized and big municipalities. Our intention was also to get an equal proportion of classes in fourth grade and eighth grade. The class sizes varied between 17 and 43 pupils.

The project was performed in 12 “experimental classes” of different schools in the middle of Sweden. All together there were 322 pupils, 193 pupils in the fifth grade (12-13 years old) and 129 pupils in the ninth grade (15-16 years old). There were 12 “control classes”, 374 pupils together, 169 pupils in the fifth grade and 205 in the ninth grade. The training time was around 2 hours a week over 12 months, from April 2002 to April 2003.

The LEGO Dacta material in the classroom work

The LEGO material, ie. the construction kit, consisted of a mechanical assembly system, a set of sensors and actuators, a central control unit (the programmable brick), a programming environment, ie. a computer and software and working instructions and manuals. The programmable brick is the most noticeable component of the kit, as it provides both control and power to all LEGO constructions.

When the experimental classes worked with the LEGO material, pupils cooperated in smaller groups, generally 3-4 pupils, which we called working groups. The groups did not follow a certain syllabus working with the LEGO material. The work was rather adjusted to the ordinary school activities and to the local preconditions at each school. To be sure that the classes started in a similar way, the teachers were instructed to start with the same kind of task.

It was stated in the research plan that the pupils should work with the LEGO material about eight hours a month. This level was decided to ensure that pupils should have similar preconditions. It was planned that the researchers should come out and visit the EG during this period, to follow up how well the work continued and what changed during this time.

We noticed that the school classes solved the practical things in different ways. The physical preconditions concerning rooms varied at the different schools. Some classes could have LEGO material out permanently, while other classes took the material out each time before using it. Some had to work in special computer rooms, other classes needed to go to other school rooms to get computers, and other classes had only a couple of computers in their ordinary class room which they used to program the robotics the pupils built.

In some classes, especially in grade 8, the teachers have been able to integrate the work in their ordinary teaching, in mathematics or in technology. This was not usually the case in grade 4, where LEGO work was more like a separate part of the school day. The participating teachers in the project have been taught how to handle the LEGO Dacta material, as pupils often ask highly intricate questions.

Methods used in the project

In this research project we used both quantitative and qualitative methods, i.e. triangulation (see for example Reason & Rowan (1985)).

Qualitative methods

The qualitative methods used in the project were observation, interview and inquiry.

Besides the above, every teacher involved in the project had to document their LEGO lessons; what kinds of activities were done during the lessons, how the pupils worked, what kind of problems they faced, and so on. The researchers made regular visits to the project schools. While visiting they made interviews with the pupils and observations of the work. The interviews were taped for later analysis. The purpose was to be able to follow the working group’s achievements during the project as well as discern their attitudes/feelings about the LEGO material. Also we could analyse their way of speaking about the material and their understanding of the LEGO concepts.

Quantitative methods

The quantitative methods consisted of different tests in mathematics and problem solving, and also covered inquiries concerning pupils' attitudes to working with robotics.

Before the school classes started their work with the LEGO Dacta material they had to perform a test in mathematics that was similar to the Swedish national test in mathematics, followed up by a test at the end of the project. The test was done by both EG and CG. The purpose was to compare the achievements for the two groups to see if there had been any changes/improvements. The problem solving test for both EG and CG at the same period was given for the same reason.

Method for studying interrelationship

The response for these pupils against the control pupils (which may depend on the MPA), can be described by a regression model with binary response variable. The usual logistic regression is used to study how the MPA are related to a dichotomous outcome Y = 0 (control pupils) or Y = 1 (training pupils). Y is the dependent variable in the logistic regression and is a function of a number of explanatory variables associated with their respective parameters. The maximum likelihood estimate of the parameters is obtained after maximizing the log-likelihood function. These estimated parameters are the change in the following logarithm of the odds ratio: $\log\left[\frac{\Pr ob(Y=1\,|\,MPA)}{1-\Pr ob(Y=1\,|\,MPA)}\right]$.

In this paper, we study issues related to the model above in a multilevel perspective, i.e. by applying the multilevel logistic regression to our data. We have a clear case of a system in which individuals are subject to the influences of grouping pupils to learn in classes. The units in such a system lie at two different levels of a hierarchy. Note that this hierarchy should be taken into account and standard models and estimation methods are not appropriate in this case and will give biased and misleading results. A typical suitable multilevel model, Goldstein (1995), of this system would assign pupils to level one and classes to level two. Units at level one are recognized as being grouped, or nested, within units at the next higher level. In this case we can see clear hierarchical structure in the data. From this point of view of our model what matters is how this structure affects the measurements of interest. Thus when measuring educational achievement, it is known that the average achievement varies from one class to another. This means that pupils within a group will be more alike, on average, than pupils from different groups.

The point of multilevel modeling is that a statistical model should explicitly recognize a hierarchical structure where one is present, and if this is not done then we need to be aware of the consequences of failing to do this, which leads to inefficient modeling. It is inefficient because it involves estimating many more coefficients than the multilevel procedure, does not treat groups as a random sample and provides no useful quantification of the variation among classes or groups in the population more generally. However, data with more than one level structure (as in our study) usually is analysed using a multilevel model. The multilevel model we use in this paper helps us investigate many important issues related to the study. Patterns of variability, in particular the nested form of variability, are studied with respect to different variables of interest. The fixed effect and the random effects are part of the outcome of the multilevel model that helps us decide the significance and direction of each effect.

Very often it makes sense to use probability models to represent the variability within and between groups, i.e. to perceive the unexplained variation within groups and the unexplained variation between groups as random variability. For a study of pupils within groups, for example, this means that not only unexplained variation between pupils, but also unexplained variation between groups or classes are regarded as random variability. This can be expressed by statistical models with so-called random coefficients. The hierarchical model is such a random coefficient model for multilevel data and is by now the main tool for multilevel analysis.

If we consider any outcome point in our study we may notice that it comes from a sampling design called a multi-stage sample. A simple random sample is selected from the first stage followed by a second stage random sample selection, nested in the first stage. Take for example the student score in mathematics. A cost-efficient strategy of evaluating schools and students within schools is by taking a random sample of schools and then a second stage random selection of student sample from each school. This method of sampling represents a two-stage design. However the mathematics score variance consists of two components, between pupils within any school variance and between schools variance. The calculations in this study are performed using the statistical program packages, MLWIN-version 1.1 and SPLUS-version 6.1.

Methods for opinion studies

Pupils in the LEGO study are classified into categories of options that represent their opinion in a given question. We call the pupils in this classification "raters". The questioner includes 17 questions with each question consisting of 4 categories (see the appendix) to investigate the training effects after the participation of the pupil in this training program. At this stage, we are mainly interested in investigating the following:

1. Testing that the distribution between the 4 categories is the same.
2. Whether the after training choice is independent from the before training choice for the pupils who changed their mind?
3. Whether there is any symmetric association pattern around the original answers?

Many of the social studies involve responses taken on the ordinal scale and may involve subjective opinion and, in particular, an intermediate category will often be subjective to more misclassification than an extreme category because there are two directions in which to err from the extremes. A square table can be used to display the joint rating of two occasions (before and after).

Tables of equal rows and columns with dependence structure have two matters. The first, important issue is the difference in the marginal distribution. The question that we like to have an answer for is "whether classification for after training tends to be higher than before training"? Secondly, the extent of main-diagonal occurrence within the joint distribution of the rating and the question one may ask is "whether agreement between the two occasions, e.g. non-significant or diagonal, leaks"? Perfect agreement occurs when the total probability of agreement = 1. Pre and post results obtained from these raters' data are analysed using Generalized Linear Model (GLM) and the data is in square table format. Matched pairs models are used to analyse agreement between two observers. Moreover, the structure of agreement and fitting different models is also investigated.

Now we consider the pupil's opinion before and after LEGO training. For any pupil, indexed by s, we suppose that the log probabilities for the k response categories are

$\lambda_{s1},\ldots,\lambda_{sk}$ Before LEGO Training, and

$\lambda_{s1}+\tau_1,\ldots,\lambda_{sk}+\tau_k$ After LEGO Training

So that $(\tau_1,\ldots,\tau_k)$ is the training effect, assumed to be the same for all individual pupils. To simplify the idea of modelling of square tables we will introduce the main type of models that we use in the analysis. The first is the symmetry model where we like to test the null hypothesis

$$H_0 \quad : \qquad \pi_{ij} = \pi_{ji},$$

that is, we like to test if the cell probability on one side of the main diagonal are a mirror image of those on the other side.

For expected frequency the logit value is

$$\log \mu_{ij} = \lambda + \lambda_i + \lambda_j + \lambda_{ij} \qquad (1)$$

The quasi symmetry model implies some association and allows the main effects terms to differ so that

$$(\log \mu_{ij} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_{ij}) \neq (\log \mu_{ji} = \lambda + \lambda_j^X + \lambda_i^Y + \lambda_{ij}) \qquad (2)$$

The main effect in (2) is different for rows and columns and the resulting estimates are the differences in $\{\lambda_j^Y - \lambda_j^X\}$ for j = 1,2,…, .

The marginal homogeneity model compares marginal distributions in the square table, and one may test the hypothesis of marginal homogeneity by comparing the fit of the specific square table. It is known, that a table that satisfies both quasi symmetry and marginal homogeneity can also satisfy symmetry and one can test marginal homogeneity by comparing goodness-of-fit statistics for the symmetry and quasi-symmetry models (Bishop et al., 1975).

Conditional symmetry models may be used when categories are ordered, and this model estimate extra parameter τ for the off diagonal on structure of agreements. The main objective of this extra parameter is to quantify the effect on the structure of agreement. The symmetry model is a special case when $\tau = 0$. The following model represents a generalization that includes the condition when symmetry does not hold with ordered category.

$$\log \mu_{ij} = \lambda + \lambda_i + \lambda_j + \lambda_{ij} + \tau I(i < j), \quad I(\bullet) \text{ is the indicator function} \qquad (3)$$

Results

In this section we present the most important qualitative and quantitative results of this study.

Qualitative results

Different strategies of learning the material

We have found that pupils learn the material in various ways, which is in line with other research (Kafai, 1996). One way to learn by children is by a “trial-and-error method”. Another way is more “cooperative”: by asking their fellow workers. Alternatively they ask another pupil in the class that is considered to know the material much better than oneself.

More seldom are situations where students read the written instructions from the teacher or from a book. One reason may be the normal situation when pupils constructed the robotics, which often seems to be a situation of “joyful stress”, but sometimes they took help from instructive pictures. We could also observe that boys were more often less willing to follow the instructions, whereas girls seemed to be more concentrated and intent on following the written task.

Pupils learning

Even though our study did not report significant contributions as far as logic skill, we could still discover other interesting things. Of certain interest was the pupils’ learning to cooperate in working groups. On interviewing pupils we often heard how the work with LEGO had enhanced the feeling of a community.

Another positive thing was the pupils’ better understanding of how to program the computers and how to load different programs to the robotics via an infrared beam. Gradually the pupils became more conscious of how they could steer the robotics by making and loading complex programs which made the robotics perform and interact in an advanced way.

We did not see any significant difference between the younger and older pupils concerning the ability to build, program and more generally handle the LEGO material. The same holds for gender, there was not any general difference between girls and boys.

Learning context

From our observations we could draw several conclusions about how the environment should be best organized to best stimulate pupils’ learning of the LEGO material. First, there needs to be a large space for the pupils to work: they must be able to spread the LEGO material on the ground, to “play around”, and test different kind of solutions for each kind of project they face. The working groups should not be too big (maximum 2-3 pupils/LEGO box). The task given to the pupils must be concrete, relevant and realistic to solve. It is very important that the pupils can relate the material to their ordinary school work and their different subjects.

The role of the teacher

The role of the teacher, as a mediator of knowledge and skills, was crucial for coping with problems related to this kind of technology (Chioccariella et al., 2001). The teacher must be able to support the pupils and help them understand the LEGO Dacta material on a deeper level. We have seen from the study the advantage of a well prepared teacher who can meet even advanced problems from eager pupils. Sometimes tricky problems had to be resolved before the pupils could continue their work with new constructions. A teacher with a background in physics or natural science could discuss LEGO technology in comparison with other technology and give a broader perspective to the available technology.

One opinion often heard among the teachers was about the advantage of two teachers working together in the classroom. This made it possible for one teacher to stop and discuss situations thoroughly with a working group of pupils, while the other one had control of the classroom situation. Teachers or other adults sometimes also played another important role, as the resolver of conflicts. Even though we generally found collaboration in working groups good, sometimes conflicts occurred. A trained supervising teacher could easily be helpful in resolving any arguments by explaining to the pupils the opportunities which open to them through cooperating together (Knoop, 2002).

Quantitative results

Comparing control and trained groups in the fifth grade

In this part of the study we start by investigating the properties of any possible differences between the control and the trained groups for the pupils in the fifth grade (ie. after the LEGO training experiment). We first start by applying the binomial second order marginal quasi-likelihood (MQL) approach (for more information about this approach, see Hox, 2002 and Snijders et al., 2002).

Our analysis reveals that the vast majority of unexplained variation in the probability of LEGO training is, as expected, between classes which support our design procedure of dividing these classes into two groups (trained and control). Moreover, we find that pupils that highly support the idea that "pupils can co-operate better in working groups" belong with high probability to the trained group. When using the multilevel regression with results from mathematics as a dependent variable, we find a common agreement among the pupils that are good at mathematics in general regarding the attitude to question 6, namely that "Working with LEGO material helps us learn things like order and method". On the other hand, we find a negative output for question 5, that "LEGO-material should not be used in the school". Note that there is a negation in this question which means in this case that the pupils do not support this statement (most of them totally disagreed with the statement). Moreover, we also find a negative output for question 12, that "Girls and boys learn equally well when working with LEGO robots" indicating that the boys and girls in fact do not believe that they learn equally well.

Table 1. Results from the multilevel regression in two stages

Fixed Effect	Model Stage 1	Standard error	Model Stage 2	Standard error
Intercept	29.08	0.476	26.5	1.683
Question 5: LEGO-material should not be used in the school.			-0.601	0.314
Question 6: Working with LEGO material helps us learn things like order and method.			0.811	0.363
Question 12: Girls and boys learn equally well when working with LEGO robots.			-0.809	0.400
Random Effect				
Level 2 variance	2.235	1.272	1.9	1.1
Level 1 variance	25.188	2.039	24.4	2.0
Log Likelihood	1961.4		1947.6	

The sample size consists of 165 pupils from different schools.

Table 1 summarises the multilevel results of the estimated parameters together with their standard errors from the fixed effect once using the empty model with only a constant term included and then using the full model with the attitude questions included (here we only find questions 5, 6 and 12 to be statistically significant effects). The second part of the table, i.e. the random effect part, shows the level one and level two variance estimates together with the log likelihood which indicate that the second model is more proper than the first. By adding the fixed effect, i.e. questions 5, 6 and 12, to the model, we expand the variation in the outcome mathematics (the difference in deviance is 13.8 (=1961.4 - 1947.6). The level one variance (24.4) with its respective standard error (2.0) shows significance in the differences between pupils when having the mathematics as the dependent variable. In other words, the vast majority of unexplained variation in mathematics comes from the difference between students. On the other hand, when looking at the level one variance (1.9) together with its standard error (1.1), we find the class effect not to be significant in this case.

When looking at achievements in mathematics for this group of pupils before and after the training by using the standard two-sample t-test, we find a positive shift in the mean from 0.711 to 0.817 with p-value = 0.000 (which means significant at all significant levels) indicating better performances in mathematics for the trained group. For the problem solving, on the other hand, we find a slight shift in the opposite direction from 0.696 to 0.649 with p-value = 0.023 which is rather significant.

When looking at the pupils in the trained group in the fifth grade evaluated by their teachers using level of engagement as a dependent variable, see Table 2, we find a clear difference between pupils and the overall level of engagement (3.37) which is above the average (2.5). On the other hand, we do not find any significant differences between the teachers from the different classes (when evaluating their pupils).

By adding the fixed effect, i.e. mathematics, to the model we expand the variation explained in the outcome engagement (the difference in deviance is 12.9 (= 472.5 - 459.6). Moreover, we find that the pupils with a greater ability in mathematics to be more engaged in the learning process, as shown in the following table:

Table 2. Results from the multilevel regression in two stages

Fixed Effect	Model Stage 1	Standard error	Model Stage 2	Standard error
Intercept	3.36	0.11	1.73	0.45
Mathematics			0.06	0.02
Random Effect				
Level 2 variance	0.04	0.05	0.02	0.03
Level 1 variance	0.99	0.11	0.94	0.10
Log Likelihood	472.5		459.6	

The sample size consists of 165 pupils from different schools.

When adding the attitude variables in the modelling, our results (not shown here) reveal that the vast majority of the pupils agree with respect to question 5 that says “LEGO material ought to be used already in pre-school”, which means the pupils that highly support this statement tend to be more engaged in the learning process. We have the same result with respect to question 7 which indicates that working with LEGO material can give the pupils more self confidence. Moreover, looking at the level two variance (0.94) with its standard error (0.10), we find significant differences between students when evaluated by their teachers using a level of engagement as the dependent variable. The level one variance (0.02) with its respective standard error (0.03) does not indicate any significance difference between teachers’ evaluations of their students in different classes using the engagement level as a dependent variable.

Comparing control and trained groups in the ninth grade

Here, we investigate the properties of any possible differences between the control and the trained groups for the pupils in the ninth grade after (i.e. after the LEGO training experiment). When applying the (MQL) approach we could not find any significant differences between the controlled and trained groups regarding the mathematics and problem solving. However, we find a common agreement on refusing the statement which states “Teaching about LEGO material does not require any computer knowledge”. In particular, pupils who highly disagree are good at mathematics.

When looking at achievements in mathematics for this group of pupils before and after the training by using the same standard two-sample t-test as for grade 5, we did not find any significant shifts in the mean with regard to mathematics or problem solving.

When looking at the pupils in trained grade 9 evaluated by their teachers using level of engagement as a dependent variable, we find a clear difference between pupils, and the overall level of engagement (3.13) which is above the average (2.5). The classes show no significant effect on the level of engagement, but as we mentioned the pupils differ in their level of engagement, see Table 3:

When including the mathematics as an explanatory variable, the result shows that this variable is significant with a positive sign indicating that pupils with high ability in mathematics tend to be more engaged. More variation is explained when including this variable in the model; the difference in the deviance is 18.2 (= 375.2 - 356.97), see Table 3 below:

Table 3. Results from the multilevel regression in two stages

Fixed Effect	Model Stage 1	Standard error	Model Stage2	Standard error
Intercept	3.13	0.22	1.80	0.39
Math			0.04	0.01
Random Effect				
Level 2 variance	0.22	0.18	0.30	0.21
Level 1 variance	1.46	0.20	1.22	0.17
Log Likelihood	375.2		356.97	

The sample size consists of 114 pupils from different schools.

As in the case of grade 5 pupils, we find the attitude questions number 5 and 7 to show the major agreement among the pupils.

Analysing the opinion studies

We here analyse the differences in responses between the two occasions, before and after. We used the method of factor analysis to select the most important questions from the list of 17, (see the Appendix).

Then, when considering the trained group in the fifth grade and based on the method of factor analysis, we find that following questions 4, 13, 16 and 17 to have the greatest factor loadings (i.e., the most important variables) both before and after training. For the trained group in the ninth grade pupils, we find questions 4, 6, 9 and 12 to be most important. Accordingly we analyse these questions, regarding the possible opinion changes among the pupils, by means of the method of GLM for matched pair models that we mentioned in the previous section. The Quasi symmetry model, however, is shown to have a better fit than the symmetry model and hence we analyse results that are obtained by using the Quasi symmetry method only (results supporting this are excluded to save space but are available upon request from the authors).

Generally, the results, regarding equations 4, 13, 16 and 17 for the fifth grade and equations 4, 6, 9 and 12 for the ninth grade, indicate that pupils with a high level opinion (i.e., those who totally agree) before training have changed their responses, regarding these questions, to a lower level (i.e. partly agree) after training. This is happening more often than the reverse. The results also indicate that pupils' option choices after training are independent from their option choices before training. Moreover, the results reveal that there is a symmetric association pattern and that any lack of symmetry seems to reflect slight marginal heterogeneity.

Concluding Discussion

There has been an intense debate in pedagogical spheres about the outcomes of LOGO/LEGO, which is mainly a part of an even broader discussion about justifying the use of computers in education. A good overview of this discussion is the paper by Harvey (2001) "Harmful to children?" where he analyses why the idea of educational computer is under attack.

Within the LEGO research group reasonable statements have been given why use of what they call the new generation of computationally-enhanced manipulative materials, or shortly "digital manipulatives". These researchers mean that the traditional field of instructional design is of little help: it focuses on strategies and materials to help teachers instruct. Instead, the interest must be in developing strategies and materials to help children construct. They call this effort "constructional design" (Resnick, 1996; Resnick, Bruckman, & Martin, 1996). Constructional design is a type of meta-design: it involves the design of new tools and activities to support children in their own design activities. In short, constructional design involves designing for designers.

A lot of research has been performed to find out whether or not there are any substantial benefits when using LEGO in educational contexts. Most of the research – which has been shown in the research background of this article – have weaknesses due to smallness and time. The size of this study is much bigger than most of the other LEGO studies, except from the initially referred Peruvian study, that in our opinion has serious shortcomings. This project is run in a complex context; schools. Certainly there is a risk for some minor bias because of the pupils' use of LEGO in their spare time, though we consider this bias as statistically negligible.

The study was built up from both a qualitative and quantitative research perspective. It was based on two stated hypotheses, which are dealt with in order below. The results show better performances in mathematics for the trained group in grade 5. When looking at achievements in mathematics for pupils in grade 9 before and after the training, we did not find any significant shifts in the mean with regards to mathematics. For the problem solving, there is no significant improvement either. This seems to be true for both grade 5 and 9. An interesting result of the study is, that pupils with higher ability in mathematics tend to be more engaged.

There is not a general positive attitude towards LEGO among the pupils. We have seen in our study that for a certain category of pupils, i.e. high-performing pupils in mathematics in grade 5, they have a positive attitude towards the LEGO material. The same is also true for grade 9. GLM for matched pair models and the Quasi symmetry model has been used in the analysis of the opinion studies among the pupils. The results reveal that, in some of the questions, the pupils have changed their opinion responses from high-level opinion before training to a lower level after training. This happened more often than the reverse.

Summarising, this field test shows a complex variation of data concerning the different measured variables. We would like to point out that lots of data also have been collected by using other methods than the quantitative ones. Thus several qualitative methods have been used, among others, including interviews and observations. Briefly, these other methods confirm our results above. The qualitative and quantitative results are in alignment, where the qualitative data has been necessarily comprehensive, and the qualitative data adds learning approaches which is not evident from the quantitative data.

After reviewing research in this area, we conclude that it is difficult to confirm the hypothesis that LEGO - generally - has positive effects on cognitive development. However our study indicates certain positive effects can be shown for groups/categories of pupils. Pedersen (1998) argues that the teacher's role in the positive results is important. It is likely that the teacher also has considerable influence over the way in which these tools are received by the pupils, boys as well as girls. In this regard, parallel arguments can surely be made for an interest in technology.

More follow-up studies, especially longitudinal, need to be conducted to answer the question of whether LEGO training which stimulates a pupil's ability to solve logical problems can really be raised by using LEGO Mindstorms for Schools? The question of the extent to which LEGO Mindstorms for Schools activities can generally raise the level of problem solving is indeed complex, and will surely need to be examined in further studies.

References

Becker, H. J. (1987). The importance of a methodology that maximizes falsifiability: Its applicability to research about Logo. *Educational Researcher, 16* (5), 11-16.

Bishop, Y. V., Fienberg, S. E., & Holland. P. W. (1975). *Discrete Multivariate Analysis*, Cambridge: MIT Press.

Chioccariello, A., Manca, M., & Sarti, L. (2001). *Children's playful learning with a robotic construction kit*, Genova: Instituto per le Tecnologie Didattiche – C.N.R.

Goldstein, H. (1995). *Multilevel Statistical models. (2nd Ed.)*, London: Edward Arnold.

Harvey, B. (2001). *Harmful to children? The Alliance for Children Report*, Berkeley,: University of California.

Hox, J. J. (2002). *Multilevel Analysis - Techniques and Applications*, Mahwah: Lawrence Erlbaum.

Iturrizaga, I. M. (2000). *Study of Educational Impact of the LEGO Dacta Materials - INFOESCUELA - MED*. Final Report, retrieved June 30, 2006, from, http://www.lego.com/education/download/infoescuela.pdf.

Kafai, Y. (1996). Learning Design by Making Games: Children's Development of Design Strategies in the Creation of a Complex Computational Artifact. In Kafai, Y. & Resnick, M. (Eds), *Constructionism in practice: Designing, thinking, and learning in a digital world*, Mahwah, New Jersey, USA: Lawrence Erlbaum.

Knoop, H. K. (2002). *Play, Learning & Creativity - Why Happy Children are better Learners*, Köpenhamn: Aschehoug Forlag.

Lave, J. (1988). *Cognition in practice*, Cambridge, UK: Cambridge University Press.

Lindh J., Carlsson, S., Kilbrink, N., & Segolsson, M. (2003). *Programmable Construction Kits in an Educational Environment*, final report, Jönköping: School of Education and Communication.

Marton, F., & Booth, S. (2000). *Om lärande (About learning)*, Lund: Studentlitteratur.

Moore, L. M., & Ray, B. K. (1999). Statistical Methods for Sensitivity and Performance Analysis in Computer Experiments. In Farrington, P. A., Nembhard, H. B., Sturrock, D. T. & Evans, G. W. (Eds.), *Proceedings of the 1999 Winter Simulation Conference*, 486-491, retrieved June 30, 2006, from, http://www.informs-cs.org/wsc99papers/068.PDF.

Papert, S. (1980). *Mindstorms - Children, Computers, and Powerful Ideas*, New York, USA: Basic Books.

Papert, S. (1993). *The Children's Machine: Rethinking School in the Age of the Computer*, New York: Basic Books.

Pea, R. D. (1985). Beyond Amplification: Using the Computer to Reorganize Mental Functioning. *Educational Researcher, 20* (4), 167-182.

Pedersen, J. (1998). *Informationstekniken i skolan: En forskningsöversikt* (98:343), Stockholm: Skolverket, Liber distribution.

Piaget, J., & Garcia, R. (1991). *Toward a logic of meanings*, Hillsdale, NJ, USA: Lawrence Erlbaum.

Reason, P., & Rowan, J. (1985). *Human inquiry - A Sourcebook of new paradigm research*, New York, USA: John Wiley & Sons.

Resnick, M. (1991). Xylophones, Hamsters, and Fireworks: The Role of Diversity in Constructionist Activities. In Harel, I. & Papert, S. (Eds.), *Constructionism: Research Reports and Essays, 1985-1990 by the Epistemology and Learning Research Group*, Norwood, NJ, USA: Ablex Publishing.

Resnick, M. (1996). Towards a Practice of Constructional Design. In Schauble, L & Glaser, R. (Eds.), *Innovations in Learning: New Environments for Education*, Mahwah, NJ: Lawrence Erlbaum.

Resnick, M., Bruckman, A., & Martin, F. (1996). Pianos Not Stereos: Creating Computational Construction Kits. *Interactions, 3* (6), 41-50.

Snijders. T. B., & Bosker, R. J. (2002). *Multilevel analysis. An introduction to basic and advanced multilevel modeling*, SAGE Publication.

Appendix A

Statement	How far do you agree with these statements?			
	Totally agree	Partly agree	Partly disagree	Totally disagree
1. LEGO material ought to be used already in pre-school.				
2. LEGO-robots are well suited for techniques lessons.				
3. LEGO material enable pupils to cooperate better in groups.				
4. Many of the pieces in the LEGO package get lost in the school.				
5. LEGO-material should not be used in the school.				
6. Working with LEGO material helps us learn things like order and method.				
7. Working with LEGO robots makes us more self confident.				
8. We learn a lot when we build and program LEGO robots.				
9. Learning about techniques helps us develop in many ways.				
10. You cannot build LEGO robots in language lessons.				
11. Boys like to compete with LEGO robots more than girls do.				
12. Girls and boys learn equally well when working with LEGO robots.				
13. LEGO robots are only toys or a game.				
14. Teaching about LEGO material does not require any computer knowledge.				
15. Boys know more about techniques than girls.				
16. Group work tends to be better when working with LEGO.				
17. We learn best by using LEGO robots in games or competitions.				