Nordic school mathematics revisited – on the introduction and functionality of New Math

JOHAN PRYTZ AND MARTIN KARLBERG

We investigate the functionality of New Math in order to nuance and deepen our understanding of the process to introduce New Math in Sweden. Our main material is data and statistics from an intervention study conducted during the preparation of the New Math reform in Sweden. In particular we investigate the claim that New Math material had negative effects on students' learning. Our conclusion is that the claim is, in most parts, incorrect.

This paper concerns the introduction of New Math in Sweden in the 1960s and 70s, in particular an intervention study[1] where new curriculum and textbook material were tested in grades 1–9. This intervention study was accounted for in the report *Nordisk skolmatematik (Nordic school mathematics)*, published in 1967.

The aim of our study reported in this article is to investigate the functionality of the New Math material in teaching. Our main material has been data from the aforementioned intervention study. In particular, we have been interested in the claim that New Math material per se had negative effects on students in general. During the 1960s and later on, critics of New Math made such claims. Thus, we are testing claims about New Math material, made in the 1960s and 70s, on the basis of data produced in the 1960s during the preparation of the New Math reform.

The rationale for doing this is linked to an on-going study of the implementation of New Math in Sweden. It is then pertinent to learn whether and in what respect the New Math material, per se, had general negative effects. If the critics were right, the questions about in what respect the reform was a failure are in part easily answered: to provide all

Johan Prytz, *Uppsala University*
Martin Karlberg, *Uppsala University*

Prytz, J. & Karlberg, M. (2016). Nordic school mathematics revisited – on the introduction and functionality of New Math. *Nordic Studies in Mathematics Education*, 21 (1), 71–93.

teachers with a really bad teaching device is in itself a failure. The answer to questions about why the implementation failed is also in part quite straightforward: people tend to abandon devices that do not work.

Short background – New Math in Sweden and abroad

The New Math can be considered an international movement during the period 1950–1970 (approx.) which was aimed at bringing about improvement of mathematics education on the basis of modern mathematics and psychology.

A basic idea of the New Math was to provide a common foundation for school mathematics using elements from modern mathematics, for instance set theory. The role model in this respect was the Bourbaki group and their attempt in the 1930s to create a foundation for scientific mathematics (Kilpatrick, 2012). But New Math was more than set theory. Important concepts were also relation and function (Bjarnadóttir, 2014). The New Math curriculum also included new topics such as logic, modern algebra, probability and statistics (Kilpatrick, 2012).

Originally, New Math was confined to secondary education; the objective was to better prepare students for university. Soon, however, the New Math ideas were expanded; leading mathematicians, mathematics educators and psychologists found them applicable throughout the school system (K–12). The idea was that a common foundation should facilitate structure, which in turn should result in better learning. This line of thinking was supported by the leading psychologists Jerome Bruner and Jean Piaget (Phillips, 2015). Bruner's standpoint was that learning would be more efficient if teaching focused on the structure of knowledge rather than the content. Both Bruner and Piaget maintained that there were similarities between mental structures and mathematical structures that could be of advantage in teaching.

Much of these ideas about psychology and mathematics concerned arithmetic. A key idea in that area was *understanding*; it should be the basis for learning rather than mechanical rote learning of facts and algorithms (Bjarnadóttir, 2014; Phillips, 2015). And understanding should be promoted if teaching focused on mathematical structures (Phillips, 2015).

But, according to Bruner, the focus on structures in one topic could also yield transfer effects in the sense that students would be able to learn other topics more efficiently. Furthermore, he did not confine this idea purely to mathematics; learning mathematics in this way could also improve thinking and learning in general (Phillips, 2015).

The New Math movement was, however, not only about mathematical and psychological theories. Also methodological issues were addressed. Concrete materials and active pedagogy were considered essential. This concerned in particular elementary teaching, where the aim was to develop understanding of concepts through a process of guided experimentation and discovery. In one of the main international reports about New Math, this method of teaching was considered more important than using the language of set theory (Bjarnadóttir, 2014).

In Sweden, New Math was introduced on a broad scale in grades 1–9 when the curriculum of 1969 took effect. The reform of the mathematics curriculum had, however, begun in the early 1960s.

An essential part in the planning process in Sweden and other Nordic countries (Finland, Denmark and Norway) was *Nordiska kommittén för matematikundervisningens modernisering (Nordic committee for the modernization of mathematics instruction)*, from here on denoted the committee or NK. The committee was established in 1960 and comprised four representatives from each of the four countries. One of the Swedish representatives, Lennart Sandgren (1926–2009), chaired the committee. Apart from these 16 members, the committee also had a so-called secretary, the Swede Matts Håstad (1931–) (NK, 1967). Here, the secretary should be understood as an executive officer. The committee also engaged about 30 people to design new curriculum documents and textbooks, but also to carry out surveys and trials. This group comprised university mathematicians, teachers from grades 1–12, school inspectors and researchers in education and psychology (NK, 1967). The main task of the committee was to design and try out new curriculum documents and texts to be used in teaching (NK, 1967). These experimental texts were tested in about 1,310 school classes in the period 1961–1966 (NK, 1967). In grades 1–6, 90 classes were involved; in grades 7–9, 450 classes; and in grades 10–12, 770 classes. However, far fewer were involved in evaluating the material.

The influence from other countries was evident. According to the final report, the committee had the Royaumont report as a starting point for its work (NK, 1967).

In an international perspective, the Royaumont conference was a key moment in the history of New Math. It is often considered the starting point of the world-wide dissemination of New Math (Bjarnadóttir, 2014; De Bock & Vanpaemel, 2015). Financially supported by the OECD, leading individuals in mathematics and mathematics education, from several countries, met for a twelve-day conference at the former Abbey of Royaumont in France. Lennart Sandgren was the Swedish representative. During the conference the discussions and efforts of the previous years were summarized and the way forward was mapped out. The conference resulted in a 246-page report entitled *New thinking in school mathematics*. The Nordic committee also based its work on previous New Math development projects (NK, 1967); for instance the School mathematics study project (USA), the Illinois project (USA), the School mathematics project

(England), but also smaller projects built around individuals like Suppes (USA), Dienes (Canada) and Papy (Belgium). In some cases the influence from the USA was even more direct; the Nordic texts tested in grades 4–6 were direct translations of texts developed by the School mathematics study group (NK, 1967).

When the final report – *Nordisk skolmatematik* – was published in 1967, the committee had been working for six years. Some classes had used the experimental texts for as long as three years. The final report includes the evaluations of the experimental texts. It is these results this paper concerns.

The explicit purpose of the evaluations was to develop the New Math material; it was not to judge whether or not the New Math material was better than traditional textbooks (NK, 1967). One reason for the latter was the lack of information about the teachers and students. Another reason was that the goals and content of the New Math curriculum were so different from the traditional curriculum; there were no common standards for an evaluation.

The latter is probably one reason why New Math was introduced in Sweden despite the fact that the evaluations did not indicate that New Math teaching was superior to regular teaching. This will be apparent later on in this paper. Another reason might be related to what was viewed as evidence. Evidently, data from teaching experiments that indicated superior learning outcomes was not considered necessary by people in power, at least not in this case. It seems as if it was general scientific arguments – however hypothetical in character – that decided the matter. As mentioned above, New Math was developed on the basis of scientific psychological results and scientific mathematics. Nonetheless, the New Math project in Sweden, as in many other countries, was the subject of criticism.

Critique about New Math having negative effects per se

Here we account for one particular type of critique of New Math: the negative effects per se critique. The reason for this is that this type of critique cannot be ignored when one is setting out to explain in what respect and why New Math reforms failed. We shall return to that issue in the next section. In this section we shall show that this line of critique existed in Sweden and that it was not an isolated Swedish phenomenon. One of the salient Swedish critics was Sven Lindström (1893–1969). He had a background as a teacher, textbook author and school headmaster. By the time of the publication of *Nordisk skolmatematik* he was an established mathematics textbook author; since 1940 he had published about 125 titles, mainly for elementary schools (grades 1–9).[2]

In 1968 he published the booklet *Överteoretisering av den elementära matematikundervisningen* (*Over-theorization of elementary mathematics instruction*) in which he put forward doubts about New Math in general and in particular the material tested by the Nordic committee.

Regarding New Math, as well as the new material and related teaching methods, Lindström concluded:

> On the basis of experience acquired, we have on the contrary reason to believe that over-theorization of teaching can have negative results, especially for learners with calculation difficulties.
>
> (Lindström, 1968, p. 72)[3]

Concerning the New Math material tested in the Swedish schools by the Nordic committee in the intervention study, Lindström had the following fears:

> First and foremost, it should be clearly understood that the planned change *cannot* be beneficial to students who lack a special aptitude for mathematics. Regarding these students – at least a substantial proportion of them – the fear is that the result of early formalization is directly negative. A comparison could be made here with the effect of forced reading instruction, which according to experts could give rise to reading and writing difficulties. It cannot be excluded that mathematics instruction that early on is focused on generalizations of mathematical processes and the use of concentrated linguistic expressions and abstract signs can give rise to inhibitions and counteract the normal development of mathematical notions. (Lindström, 1968, p. 66)

Both passages quoted are about teaching, teaching materials, teaching methods, students and students' learning; the feared problems concern precisely those entities. Nothing is said about, for instance, the implementation process and that it might be difficult to train all teachers properly. This is what we mean by the New Math material having negative effects per se.

The passages quoted above also illustrate that New Math would have negative effects in general, that is a clear majority of the students would be affected negatively. It was not only the students with most difficulties in mathematics that were in danger, but a "substantial proportion" of the students. What was meant by "substantial proportion" is not obvious, but we think it is fair to say that he meant a clear majority. In the very last sentence of the booklet Lindström wrote that he spoke on behalf of the "ordinary" student (Lindström, 1968).

Lindström's critique of the New Math contained in essence these two parts: the materials' negative effects per se and the generality of the effects. He did mention potential problems in the implementation process, more precisely the competence of the teachers and the supply of teaching material, but only briefly and on one occasion only (Lindström, 1968).

The critique regarding New Math having negative effects per se was not unique for Sweden. One of the most significant critics in USA, also known internationally, was Morris Kline (1908–1992). He published several works on the subject, but his major work in this respect was *Why Johnny can't add: the failure of the New Math*, published in 1973. Just like Lindström, he often returned to the idea that the New Math material had negative effects per se. After writing about New Math and its rigor, language, purpose and content, he concluded:

> This mathematics [the New Math] offers an abstract, rigorous version that conceals the rich and fruitful essence and emphasizes uninspiring generalities isolated from all other bodies of knowledge. It stresses final sophisticated versions of simple ideas while treating superficially the deeper ones – and so necessarily assumes a dogmatic character. The formalism of this curriculum can lead only to an erosion of the vitality of mathematics and to authoritarianism in teaching, the rote learning of new routines far more useless than the traditional routines. In brief, it presents form at the expense of substance and substance without pedagogy. (Kline, 1973, p. 102)

This quote is typical of a great deal of Kline's critique. The negative effects of New Math are due to the selection of content and the way the content is presented, that is to say the New Math material had negative effects per se. Elsewhere in the book Kline did in fact mention the implementation process, but only briefly.

Another typical feature in the quote above is the general nature of the critique; there is no reservation regarding, for instance, students' abilities, future studies or work. New Math is bad for all types of students.

Now, let us assume that Lindström and Kline were right. Then it must be a significant element in any analysis of why and to what extent a New Math reform failed. First of all, simply allowing teachers to use a bad teaching device is a failure. Secondly, people tend to abandon devices that do not function properly.

The alleged failure of New Math and explanations thereof

The subtitle of Kline's work – *The failure of the New Math* – indicates that New Math did fail or was about to fail. There are also researchers in mathematics education who have echoed these views:

> There was a world-wide movement to introduce a "new" mathematics – but we know that the effect of these efforts was negligible: Little has changed in classrooms and the change that has occurred bears little relationship to the goals of the original reform movement.
>
> (Damerow & Westbury, 1984, p. 22)

Nonetheless, whether it was a failure or not does of course depend on what is meant by New Math and failure.

One important aspect of this issue is that the New Math reform not only concerned set theory; it also brought in new topics. In Sweden this meant the introduction of new topics in earlier years, i.e. algebra and geometry, and new topics in all grades, i.e. statistics and probability. These topics have remained in all succeeding Swedish curricula (Prytz, 2015). Kilpatrick (2012) reports similar development in other countries. On a general level this is not a failure. How the content of these topics changed over time is, of course, another question and in the Swedish case this remains to be investigated.

Another aspect of reforms is that they can bring new attitudes to a school subject. Indeed there are such examples concerning New Math reforms, and changes that have remained. Kilpatrick (2012) points out that the idea about one type of mathematics course, regardless of the students' later school choices, was embraced by people within the New Math movement. If we consider the Swedish case, this idea seems to have survived even though it was first finally realized throughout grades 1–9 in 1994.

However, if we consider the Swedish case, on a policy level, and the idea of having set theory as a common foundation of school mathematics, then New Math can be considered a failure. Our argument for this is that formulations about set theory disappeared from the national curriculum of 1980. And the central position of set theory was also replaced by problem solving. Not only did problem solving become a specific topic, it was also to be an integral part of all other subjects (Prytz, 2015). Moreover, disappearance from the national curriculum was essential since the school system was centralized. All teachers and textbook authors were supposed to comply with such directives.

If we leave the policy level and consider textbooks and the actual teaching, the question about abandoning set theory is more difficult. Applying set theory was not merely a matter of using symbols and expressions from set theory; the main idea was to use ideas or concepts from set theory, which is mentioned above. This opens up the possibility that basic components of New Math did survive, but with different representations. So far, though, there are no studies on how Swedish textbooks (grades 1–9) developed in this respect. But if we consider the US case, Phillips (2015) points out that changes in textbook design that occurred in connection with the New Math reform remained well after the 1970s.

If we consider the achievements of the students, the question about the failure of New Math becomes perhaps even more complex. Kristiansson (1979) compares the results of the Swedish national test (standardprov) taken by students who followed the old curriculum and students who followed the New Math curriculum. Kristiansson (1979) uses the tests in grades 3 and 6 that were given in 1973. In general the students following the old curriculum performed significantly better. However, this is not surprising since the study was made close to the launch of the new curriculum; teachers and students may have performed better as they got used to the new curriculum. This hypothesis is confirmed by results presented by Westin (1999). He compares the national test (standardprov) given in grade 9 during the period 1973–1977. In order to get a better comparison, his study is restricted to a number of items that appeared in every test. Westin (1999) shows a clear improvement in the results during this period. But what is then meant by getting used to a new curriculum? Perhaps the teachers abandoned the key elements of the new curriculum and returned to old habits. Or they might have understood the New Math better after some years, which improved their teaching.

Thus, in the wake of New Math reforms, in Sweden and elsewhere, we can discern trails of success as well as failure, but also cases that so far are difficult to evaluate. We will now turn to explanations of successes and failures of the New Math reform.

If we look outside the Swedish context, a common approach to the issue is to consider different parts of an educational system and how they worked, or rather did not work, together; see for instance Westbury (1980), Davison & Mitchell (2008) and Amit & Fried (2008). They all observe that different groups (university mathematicians, teachers, administrators, etc.) had different views of society, science, education and school mathematics, which in turn may explain why the New Math reforms, in some respect, failed or succeeded.

Phillips (2015) has a partly different approach, as he describes how New Math in the USA fell into general disrepute in the 1970s. He places New Math in a broader societal context and argues that society, when New Math was created, was in many ways different from the society of the 1970s (Phillips, 2015). For example, in the 1950s New Math was considered an optimal means to train scientific and rational thinking, and an educational need in that respect was recognized. Moreover, science and scientists, but also authorities in general, enjoyed great public confidence. In the 1970s, public confidence in science and scientific thinking had decreased. Also the main educational needs were perceived differently.

A third type of approach is found in treatises on New Math in Sweden, in which the functionality of New Math is addressed. Kilborn, Lundberg, Selander and Öhlund (1977) study how the New Math curriculum of 1969 was prepared. They do not address explicitly why the reform was failing, but their main conclusion contains clues:

> Much of the [curriculum] text is either totally bland or too unrealistic. [...] It does not give the teacher any basis for planning the teaching. On a number of points, it is based on an unsustainable methodology whose functionality has not been tested before it was launched. (Kilborn et al., 1977, p. 1)

The two first sentences can be read as if it was the design of the curriculum text that was the problem. However, the last sentence of the quote is about New Math methodology. This suggests that the problem was related to New Math per se, but other interpretations are possible. The last sentence says that on a number of points the curriculum text is based on unsustainable methodology, which implies that on a number of other points it is based on sustainable methodology. We can understand this as suggesting that there were some sound ideas, but the curriculum text did not convey them well.

Unenge (1978, 1999) is clearer about the nature of the problem:

> As the [Swedish] program progressed for a few years, problems began to appear. While one of the major aims of the new mathematics' movement was to teach understanding, textbook presentations were rather theoretical and confused instead of clarifying concepts in student's minds. (Unenge, 1978, p. 57)

It was the textbooks (i.e. New Math materiel per se) that had negative effects on learning.

It should be mentioned that Unenge (1999) indicates that there were shortcomings in the implementation process (Unenge, 1999). Still, he is quite clear about what the main problem was.

> So it took some time before the worst madness of *the New Math* was corrected. But basically I think that other, much greater damage was done. The teachers, the good workers in the vineyards, of course realized quickly that this was madness. But it was sanctioned by the authorities, training had been organized, albeit ridiculously little, and there had even been "parent education" through the media. Confidence in the school authorities had received a blow and it is

> easy to understand that since then teachers have received news from the supreme powers with distrust – "Are new set theories on the way?" (Unenge 1999, pp. 63–64)

The main problem was the New Math, which is the object associated with madness in the quote above. And New Math was first and foremost about content and teaching methods when Unenge (1999) writes about it. The madness was not primarily related to the way the reform was implemented. Thus, Unenge is here indicating that the New Math material per se had negative effects.

Unenge (1999) also suggests that the New Math was particularly troublesome for students who found mathematics difficult.

Both Kilborn et al. (1977) and Unenge (1978, 1999) touch on an important aspect: the functionality of New Math. This aspect is not addressed by Westbury (1980), Davison & Mitchell (2008) and Amit & Fried (2008). Thus they actually overlook a possible cause of why New Math failed.

Phillips (2015) addresses this issue, but very briefly and without further analysis of the sources. He mentions that in the USA test scores in general were declining in the 1960s and 70s. Thus there were other possible causes than the New Math material of why test scores in mathematics declined. He also mentions that there were reports regarding negative as well as positive effects of New Math.

Regarding Unenge (1978, 1999) and perhaps also Kilborn et al. (1977), in our view they put too much emphasis on New Math having negative effects per se. Our analysis of the evaluation of the trials of New Math material in Sweden contradicts or gives no support to such claims. But before presenting this analysis, we shall turn to the methodological considerations behind our analysis.

Methodological considerations

In our study we have used statistical data from the intervention study reported in *Nordisk skolmatematik*. Further data have been collected from the special reports regarding the intervention study where more data and statistical details were accounted for. These reports were never published, but can be found in the Swedish National Archives (RA A, U 1369, 1371, 1425, 1431).[4]

The intervention study was designed with experimental groups and control groups and comprised students in grades 1–9 of Grundskolan and grades 7–9 of Realskolan.[5] The experimental groups received teaching with the New Math material for two or three years following a New Math curriculum.[6] The control groups received teaching with regular textbooks following the regular curriculum. More details about the number of students are provided in later sections.

The data were generated by means of tests. In grades 3, 6 and 8, the students in the experimental and control groups took the national mathematics test (Standardprov i matematik). Each group comprised about 100 students.

Comparisons of the results from the experimental and the control groups were, and still are, problematic. The New Math curriculum differed from the regular curriculum on several points. Thus, the national test suits the latter much better. Moreover, there was very little background information about the students and their teachers. Thus, it is not possible to evaluate from the results which textbooks and which curricula were most effective. All these points were underscored in the report *Nordisk skolmatematik* (NK, 1967).

However, a fourth test was done in grade 9. This time, the experimental groups were matched socially and intellectually with control groups. This test had fewer items than the national tests, but it was constructed in such a way that no group would be favoured in certain ways; the items represented parts that were common to both the New Math curriculum and the regular curriculum. Thus, this fourth test gives some indication about efficacy.

The final report and the special reports did not reveal how the actual teaching was carried out except that the experimental texts had been used. The teachers were given great freedom in carrying out the teaching (NK, 1967).

The reports did not actually contain any explicit statement about the experimental texts wholly replacing traditional textbooks. But several formulations about text series, alternative courses and alternative curriculum give strong indications that this was the case; in order to give an alternative course, you need a comprehensive alternative text material. The experimental texts also covered a vast amount of content, which indicates that they replaced the traditional textbooks (NK, 1967). Moreover, the final report provided a bit more information about the texts intended for grades 4–6. They comprised in total about 1,200 pages, which is wholly sufficient for replacing regular textbooks.

However, despite the lack of information about the teaching, we can reach some conclusions about the functionality of the New Math material.

The New Math material must be considered to have played an essential part in the teaching. Since the teachers were to teach a new kind of content, the New Math material must have been a major source of new concepts, symbols, explanations and exercises. In the reports there were no comments about teachers ignoring the experimental texts. This is also the impression we get when browsing through the vast numbers of reports from the teachers to the committee regarding their use of the experimental material (RA B). It is therefore not likely that the New Math material was often ignored and that the teachers relied solely on established traditions. Moreover, since the teachers had volunteered (NK, 1967), we can assume that they were motivated, keen to use the new material and prepared to do so. In the final report it is also assumed that there was a positive selection of teachers in terms of skill and experience (NK, 1967).

Thus, comparisons of the test results of the experimental groups and the control groups do reflect the functionality of the New Math material.

The comparative study by Kristiansson (1979), mentioned above, is for instance less suitable to evaluate the functionality of the New Math material. It is not possible to determine to what extent the teachers were motivated to follow the new course, relied on old textbooks or were not sufficiently prepared.

Regarding the selection of teachers, it adds an important dimension to our conclusions about the functionality of the New Math material. Nevertheless, it does not contradict the conclusions about it being possible to use the material successfully in the classroom. But the material might have required a skilled, experienced and motivated teacher.

The purpose of the original evaluation study was to provide data for further development of the teaching material (NK, 1967). Due to the circumstance mentioned above, the authors did not claim general validity for their results. Nor do we wish to generalize any results, but we do want to test a certain general hypothesis.

It is possible to use the data to test the general hypothesis about the New Math material having negative effects per se on all students. If the hypothesis is true, this would be reflected in the data. Thus we can confirm or reject the general hypothesis on the basis of the data at hand. For instance, if the material was completely unsuitable for teaching, the results of the experimental groups would be significantly lower on a clear majority of the test items.

Our analysis of the results of the intervention study

In grades 1–3 and grades 4–6 the students took the national test in mathematics for grades 3 and 6. This was done during the trials of the tests. The control groups in this case were the students who took the test during the trials of the test. The differences in solution frequencies (percentage units) for each item were calculated (experimental group minus control group) and each item was put in one of five categories:

++ the difference > 10 percentage units

\+ the difference in the range 3 – 10 percentage units

0 the difference in the range of -2 – 2 percentage units

\- the difference was in the range of -3 – -10 percentage units

-- the difference < -10 percentage units

The results of the comparisons are summarized in tables 1 and 2. Note that the italicized parts are our additions. The purpose is to clarify how the students performed.

We cannot give exact numbers of students in the control groups since none of the reports contained such information. The reports say only that the statistics are from the trials of the test. However, we have found the original data from the trials for the national test given in the spring of 1967. Each part was tried with a different number of students, ranging from 84 to 261 (RA C). We have no reason to believe that a sudden change in the number of students occurred from the year before. Henricson (1987) summarizes the enterprise concerning the national tests during the period 1965–1985 and does not indicate a sudden shift in the preparation of the test from 1966 to 1967.

Our conclusion is that tables 1 and 2 show that there were teachers who could use the experimental texts and achieve acceptable results. The fact that the experimental students performed really badly in about 25 percent of the items is far from the New Math material being a failure. Especially since the tests were designed for a regular course. We should also keep in mind that switching to a completely new type of teaching material may not have been an easy task for the teachers. Thus it was possible to use New Math material in teaching and attain acceptable results.

We know that all classes in grades 1–6 were undifferentiated, comprising high as well as low performing students (NK, 1967). Thus, the results of the experimental classes are not due to the absence of the less talented students in the investigated classes.

The use of the plus and minus tables may hide very large differences. When it comes to table 1 we know this for a fact, since solution frequencies for each item are accounted for in the special report on the test in grade 3 (RA A, U 1431).[7] However, these concealing effects do not contradict our conclusion. The share of items outside the double minus group is the same.

A national test was also given to students in grade 8, but this test was not taken during the trials of the national test. Another difference was that the students did not belong to undifferentiated classes. The students

Table 1. *Differences in solution frequencies, national test grade 3, study year 1965–66*

	++	+	0	-	--	*Sum*	*Share*
Mechanical calculation	2	1	2	4	5	*14*	*29 %*
Mental arithmetic	1	3	5	7	4	*20*	*41 %*
Applications	5	1	3	3	4	*16*	*33 %*
Sum	*8*	*5*	*10*	*14*	*13*	*50*	
Share	*16 %*	*10 %*	*20 %*	*29 %*	*26 %*		

Note. Eight classes were included in the experimental group. Six classes had used experimental texts for three years, two for one year. Number of experimental students: 123. Number of control students: 80–250.

Table 2. *Differences in solution frequencies, national test grade 6, spring 1966*

	++	+	0	-	--	*Sum*	*Share*
Mechanical calculation	3	3	3	4	8	*21*	*31 %*
Rough estimate calculation	1	1	3	5	8	*18*	*27 %*
Applied calculation	4	6	7	1	0	*18*	*27 %*
Geometry	5	1	2	1	1	*10*	*15 %*
Sum	*13*	*11*	*15*	*11*	*17*	*67*	
Share	*20 %*	*16 %*	*22 %*	*16 %*	*25 %*		

Note. Six classes were included in the experimental group. They had all used the experimental texts for three years. Number of experimental students: 120. Number of control students: 80–250.

attended two different types of school: Grundskolan and Realskolan. In the case of Grundskolan, there were two types of courses in mathematics: advanced and normal. All Grundskolan students in the study followed the advanced course (NK, 1967). Realskolan, on the other hand, was voluntary and the students had been selected on the basis of grades. Thus, we may assume that the proportion of students with difficulties in mathematics who took the test in grade 8 was much smaller in comparison to corresponding proportions of the students who took the test in grades 3 and 6.

Table 3. *Differences in solution frequencies, national test grade 8 (Realskolan), study year 1964–65*

	++	+	0	-	- -	*Sum*	*Share*
Equations & differences	5	2	2	4	3	*16*	*27%*
Mental arithmetic	3	2	1	3	1	*10*	*17%*
Graphic representation	4	6	1	1	2	*14*	*23%*
Applied calculation	2	5	0	4	3	*14*	*23%*
Geometry	1	0	1	0	4	*6*	*10%*
Sum	*15*	*15*	*5*	*12*	*13*	*60*	
Share	*25%*	*25%*	*8%*	*20%*	*22%*		

Note. Number of experimental students: 155. Number of control students: 99. The report does not state explicitly how long the students in table 3 had used the New Math material, but it appears to have been two years.

Table 4. *Differences in solution frequencies, national test grade 8 (Grundskolan), study year 1964–65*

	++	+	0	-	- -	*Sum*	*Share*
Equations & differences	5	0	2	5	4	*16*	*27%*
Mental arithmetic	0	1	5	3	1	*10*	*17%*
Graphic representation	1	8	3	2	0	*14*	*23%*
Applied calculation	0	1	5	4	4	*14*	*23%*
Geometry	1	0	3	0	2	*6*	*10%*
Sum	*7*	*10*	*18*	*14*	*11*	*60*	
Share	*12%*	*17%*	*30%*	*23%*	*18%*		

Note. Number of experimental students: 142. Number of control students: 163. The report does not state explicitly how long the students in table 4 had used the New Math material, but it appears to have been two years

As mentioned above, we may assume that the proportion of students with difficulties in mathematics was much smaller in the classes who took the test in grade 8 than in grades 3 and 6. However, the smaller proportion of double minus items in tables 3 and 4, in comparison to tables 1 and 2, does not suggest that New Math material had extra negative effects on students with difficulties in mathematics. First of all, the relatively small proportion of double minus items in table 4 is probably linked to the relatively large proportion of zero items in table 4. The latter was explained by some items having very low solution frequencies in both groups (NK, 1967). Observe that if the solution frequency of an item in the control group is below 10 percent, the item cannot appear in the double minus columns. Thus, table 4 is not reliable when comparing double minus items. If we then just consider table 3 with tables 1 and 2, the difference in the proportion of double minus items is less than 5 percentage units. That corresponds to two to three items. This lends no obvious support to the claim that New Math was much more difficult for students with difficulties in mathematics.

Apart from this, our conclusions and arguments regarding table 3 are the same as for tables 1 and 2: it was possible to use New Math material in teaching and reach acceptable results.

Our conclusion regarding table 4 is the same even though the numbers concerning the double minus items are unreliable. Table 4 seems to conceal a bad performance of the experimental groups, which used the New Math material. But also the control groups performed badly. Thus, the New Math material cannot have been the decisive cause of the bad performance of the experimental groups.

A second type of study was conducted in grade 9 and included data concerning students and teachers. Experimental and control groups were matched against each other with respect to school type, grades, location and the teachers' competence in mathematics.

The test included a total of twelve items in the areas of arithmetic, algebra and geometry. It was designed to be course neutral; the items were chosen from the intersection part of the New Math course and the regular course.

The results of the test are summarized in tables 5 and 6 below. The numbers are taken from the report *Jämförande matematikprov i årskurs 9* (*Comparative tests in mathematics in grade 9*), except for the ones in the columns effect size, which have been calculated by us on the basis of the numbers in the other columns. The formula is $es = (m_{exp} - m_{cont})/sd_{cont}$.

According to the report, the results were tested for significance, but it is not clear how it was done. We conducted analyses with independent T-tests and they confirm the authors' claims about significance. Note that significance says little about the size of the difference. Therefore, we have calculated the effect size.

Table 5. *Results of mathematics test with 12 items, all students: mean values, standard deviations to the mean values, number of students and effect size*

		m	*sd*	*N*	*es*
Realskola	exp.	7.48	2.30	147	-0.49
	cont.	8.55	2.18	139	
Grundskola	exp.	6.26	2.64	199	0.08
	cont.	6.00	3.13	205	

Table 6. *Results of mathematics test with 12 items, students born in 1949: mean values, standard deviations to the mean values, number of students and effect size*

		m	*sd*	*n*	*es*
Realskola	exp.	7.87	2.14	117	-0.35
	cont.	8.61	2.12	107	
Grundskola	exp.	6.22	2.67	172	0.09
	cont.	5.94	3.22	165	

However, the composition of the Realskolan and Grundskolan groups differed, which affected the numbers in table 5. The Realskolan group contained a larger proportion of older students (RA A, U 1369). If we just consider the group of students born in 1949, the results are somewhat different.

The mean value difference in the Realskolan result is still significant, but the effect size is smaller than in table 5.

For the proponents of New Math, the results in tables 5 and 6 must have been a disappointment, especially the Realskolan results. If we compare the results to today's international standards regarding interventions and effect size, they were not promising. Hattie (2009) claims that almost any intervention in school has a positive effect on academic achievement. He shows, by synthesizing more than 800 meta-analyses, that the average effect size for interventions in schools is 0.40. An effect size above 0.40 makes it possible for us to register a "real-world change" and that figure is used as a benchmark. Interventions or reforms resulting in an effect size above 0.40 should be considered worth having. An effect size below 0.40 indicates that the intervention or reform should be "regarded as in need of more consideration" (Hattie, 2009). In this perspective, the positive effect size values in the Grundskolan results in tables 5 and 6 are very small. And the Realskolan results are of course terrible.

On the other hand, we have to remember that tables 5 and 6 were based on a limited study, both with respect to mathematical content and number of students. Moreover, the study did not measure long-lasting effects. According to the proponents of New Math, the merit of New Math was that it would lead to more solid knowledge and presumably not easily forgotten knowledge.

Both tables 5 and 6 indicate that the Realskolan students were superior to the Grundskolan students in mathematics. This applies to both the control and experimental students, but particularly the control students. Moreover, the Realskolan students in the study had better grades than the Grundskolan students (RA A, U 1369). The authors of the report also claimed that the schools the Realskolan students went to had relatively high admission scores (RA A, U 1369). These circumstances are interesting in relation to our study.

In view of tables 5 and 6, the claim about the New Math material having negative effects per se on all students is contradicted by the results of the Grundskolan students. We observe virtually no effect at all. Moreover, if the claim was true, it is likely that the best students would have been least affected in a negative direction. The difference in results between the experimental groups and the control groups would then have been greatest among the Grundskolan students. In tables 5 and 6 we observe quite the opposite.

Tables 5 and 6 indicate that it was the better students (most talented or most motivated, not necessarily the same thing) who were most negatively affected by the New Math material. A possible explanation for this is that in New Math courses much emphasis was put on understanding and insight, at the expense of technical skills, e.g. more advanced computational skills. Indeed, one of the ideas behind using set theory in teaching was that it should promote understanding.

This priority was also declared in the final report (NK, 1967). From other parts of the final report we can understand that the work done to get the students to understand took time in the experimental classes. The teachers sent in reports about their teaching and from their accounts we can discern that time was spent on reading texts, at times difficult texts, and on discussions of concepts and symbols (NK, 1967).

If to every new concept or topic there were attached longer texts and discussions, involving set theory, then there was probably less time to practise technical skills in comparison to the regular course. The regular teaching would then have offered more time to spend on exercises, which would have benefited more talented students; they did not need detailed explanations and discussions in order to understand, they did so anyway. Instead they could have encountered more exercises, but also more complicated exercises. When the time in the New Math course was dedicated to other matters, the talented students simply did not get as far as on a traditional course.

It is also possible that more motivated students, that is to say more competitive students, benefited from a traditional course where more time was spent on exercises instead of reading and discussions. It might be that progress seemed clearer when a student could excel in the number of completed exercises.

Conclusions

Our analysis shows that it was possible to use the New Math material and attain acceptable results. More precisely, the students in the experimental classes performed at a level with the students who received regular instruction. However, we cannot dismiss the fact that the material needed an experienced teacher with a special interest in New Math. Our analysis also contradicts the claim that New Math material had negative effects per se on students in general. Moreover, we find no support for the claim that especially students who found mathematics difficult were disadvantaged by the New Math material. However, more talented or motivated students seem to have been hampered by the new curriculum and the new textbook material.

With these conclusions we contradict the explanations of why New Math failed in Sweden offered by Kilborn et al. (1977) and Unenge (1978, 1999). They imply more or less explicitly that the New Math material was a bad teaching device, which in turn should explain the failure.

Explanations for corresponding failures in other countries, see for instance Westbury (1980), Davison & Mitchell (2008) and Amit & Fried (2008), concern the implementation process and how the educational system functioned. On the basis of our conclusions we believe such an explanation is needed also for the Swedish case. This will be a main topic of a coming paper.

Sources

Riksarkivet (RA A) (Swedish National Archives)
Arkiv: Nordiska kommittén för modernisering av matematikundervisningen (Archive: Nordic Committee for modernization of mathematics teaching)
B1:4, Utgående skrivelser nr 1126–1497 (Outgoing letters)

Dokument:
U 1369 Jämförande matematikprov i årskurs 9 (Comparative tests in mathematics in grade 9)
U 1371 Standardproven i matematik använda i försöksklasser i årskurs 8 (Standardized tests in mathematics used in the experimental classes in grade 8)
U 1425 Prov i årskurs 6 (Test in grade 6)
U 1431 Standardprov i årskurs 3 (Standardized tests in grade 3)

Riksarkivet B (RA B)
Arkiv: Nordiska kommittén för modernisering av matematikundervisningen (Archive: Nordic Committee for modernization of mathematics teaching)
F1:11–16 Lärares rapporter om försöksverksamheten till kommittén, 1964–1967 (Teachers' reports to the Committee on the trials, 1964-1967)

Riksarkivet C (RA C)
Arkiv: Skolöverstyrelsen (Archive: National Board of Education)
Sektionen för prov och utvärdering, L4:1, L4:3 samt programmet för prov 1982–1984 (Section for test and evaluation, L4: 1, L 4: 3 and the program for tests 1982-1984)
HIad:2, Data och statistik från utprövningen av standardprov i matematik för åk 6, våren 1967 (Data and statistics from the testing of the standardized tests in mathematics for grade 6, Spring 1967)

References

Amit, M. & Fried, M. N. (2008). The Complexities of change: aspects of reform and reform research in mathematics education. In L. English (Ed.), *Handbook of international research in mathematics education* (2nd Edition) (pp. 385–414). New York: Routledge

Bjarnadóttir, K. (2014), History of teaching arithmetic. In A. Karp & G. Schubring (Eds.), *Handbook on the history of mathematics education* (pp. 431–458). New York: Springer.

Damerow, P. & Westbury, I. (1984). Conclusions drawn from the experiences of the new mathematics movement. In P. Damerow, M. E. Dunkley, B. F. Nebres & B. Werry (Eds.), *Mathematics for all: problems of cultural selectivity and unequal distribution of mathematical education and future perspectives on mathematics teaching for the majority* (Science and Technology Education Document Series No. 20, pp. 22–25). Paris: UNESCO.

Davison, D. M. & Mitchell, J. E. (2008), How is mathematics education philosophy reflected in the math wars? *The Mathematics Enthusiast*, 5(1), article 15.

De Bock, D. & Vanpaemel, G. (2015), Modern mathematics at the 1959 OEEC seminar at Royaumont. In K. Bjarnadóttir, F. Furinghetti, J. Prytz & G. Schubring (Eds.), *"Dig where you stand" 3. Proceedings of the third International Conference on the History of Mathematics Education* (pp. 151–168). Uppsala University.

Hattie, J. (2009). *Visible learning: a synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.

Henricson, S. (1987). *Skolöverstyrelsens provverksamhet: en översikt 1965–1985* (The testing enterprise of the National Board of Education: a survey 1965-1985). Stockholm: Skolöverstyrelsen.

Kilborn, W., Lundberg, S., Selander, S. & Öhlund, U. (1977). *Hej Läroplan!: hur man bestämmer vad våra barn ska lära sig i matematik* (Hello Curriculum!: how to decide what our children should learn in mathematics). Mölndal: Pedagogiska institutionen, Göteborgs universitet.

Kilpatrick, J. (2012). The new math as an international phenomenon. *ZDM: the international journal on mathematics education*, 44(4), 563–571.

Kline, M. (1973). *Why Johnny can't add: the failure of the new math*. New York: St. Martin's Press.

Kristiansson, M. (1979). *Matematikkunskaper Lgr 62, Lgr 69* (Knowledge in mathematics in the curricula of 1962 and 1969) (Ph. D thesis). Göteborgs universitet.

Lindström, S. (1968). *Överteoretisering av den elementära matematik-undervisningen: en kritisk granskning av förslaget ill studieplan i matematik för grundskolans låg- och mellanstadium* (Over-theorization of elementary mathematics instruction: a critical review of the proposed mathematics syllabus for primary school grade 1–3 and 4–6). Stockholm: Almqvist & Wiksell.

Nordiska kommittén för modernisering av matematikundervisningen (1967). *Nordisk skolmatematik* (Nordic school mathematics). Stockholm: Nordiska rådet.

Phillips, C. J. (2015). *The new math: a political history*. Chicago: The University of Chicago Press.

Prytz, J. (2015), Swedish mathematics curricula, 1850–2014. An overview. In K. Bjarnadóttir, F. Furinghetti, J. Prytz & G. Schubring (Eds.), *"Dig where you stand" 3. Proceedings of the third International Conference on the History of Mathematics Education* (pp. 309–326). Uppsala University.

Unenge, J. (1978). Mathematics in an undifferentiated school system. In F. Swetz (Ed.), *Socialist mathematics eucation* (pp. 209–255). Southampton: Burgundy Press.

Unenge, J. (1999). *Skolmatematiken i går, i dag och i morgon – med mina ögon sett* (School mathematics: yesterday, today and tomorrow – in my opinion). Stockholm: Natur och kultur.

Westbury, I. (1980). *Change and stability in the curriculum: an overview of the questions* (Research reports from the Curiculum Laboratory, College of Education, University of Ilinois at Urbana-Champaign, Number 6). Retrieved from http://files.eric.ed.gov/fulltext/ED196706.pdf

Westin, H. (1999). *Farväl standardprov: standardproven i matematik 1973–1997 för åk 9: jämförelse av resultat på uppgifter som återkommit genom åren* (Farewell standardized testing: standardized tests in mathematics 1973–1997 for grade 9: comparison of results on items reoccurring over the years). Stockholm: Lärarhögskolan, PRIM-gruppen.

Notes

1 What is meant by intervention study in this case is explained in the section Methodological considerations.

2 The numbers are from the Swedish national library database – LIBRIS.

3 A number of quoted passages appear in the paper. Most of them are from texts in Swedish. The translations have been made by the authors of this paper in cooperation with a professional translator.

4 All reports were authored by Matts Håstad, in some cases in cooperation with Erland Albihn.

5 Before 1962 there was a parallel school system in Sweden. The two main types of school were Folkskolan (1–7, primary schools) and Läroverken (5–12, secondary schools). Between 1905 and 1962, the lower part of Läroverken (5–9) was named Realskolan. By the 1962 curriculum, Grundskolan (1–9) replaced Folkskolan and Realskolan.

6 The texts used in the experimental groups were authored by the following people: Johan Amundsson (Geometry, 7–9), Gunnar Bergendal (Geometry, 7–9), Carin Claesson (all topics, 1–3), Ove Hemer (Geometry, 7–9), Matts Håstad (all topics, 1–3, Algebra, 7–9), Margareta Kristiansson (all topics, 1–3), Bertil Nyman (Geometry, 7–9) and Nils Sander (Geometry, 7–9). The text series used in grades 4–6 was a literal translation into Swedish of a US textbook series (nine books) produced by the School Mathematics Study Group. No further bibliographical data was given.

7 Regarding table 1 and the results in grade 3, the average difference for the double minus items was -20 percentage units. It is also here (in comparison to the double plus category) we find the greatest differences; the greatest

and second greatest negative differences were -55 and -46 percentage units; the third greatest was -23 percentage units. The corresponding average for the double plus items was 15 percentage units. The two greatest positive differences were about 23 percentage units.

Johan Prytz

Johan Prytz is senior lecturer at Uppsala University, Department of Education. He is mainly doing research on the history of mathematics education. Other fields of interest are textbook research and the role of language in school mathematics.

johan.prytz@edu.uu.se

Martin Karlberg

Martin Karlberg is senior lecturer at Uppsala University, Department of Education. His research is primarily focused on classroom management, behavior management and inclusive education. He specializes mainly in quantitative method.

martin.karlberg@edu.uu.se