Collaborative data teams for school improvement

Working paper, please do not cite without permission

Kim Schildkamp[1] & Adam Handelzalts

Abstract
Data-driven decision making is becoming more and more important in education. The use of data can lead to school improvement (e.g. increased student achievement). However, in several schools data are not being used. In this paper we describe the first results of a pilot into supporting schools in the use of data by a data team approach. Collaborative data teams are teams of teachers (4-6) and (assistant) school leaders (1-2), who collaboratively use data to solve a certain educational problem within the school using a structured approach. A data team defines a problem, develops hypotheses about what causes the problem, collects data to either accept or reject the hypotheses, analyses and interprets the collected data, comes up with conclusions and measures to solve the problem, and evaluates the effectiveness of the measures. The results of the pilot show that data teams can be effective in two different manners: (1) a data team uses data to solve a problem, thereby improving education and (2) data team members learn how to use data to improve their work (e.g. professional development).

Keywords: data use, data teams, data characteristics, school organizational characteristics, team characteristics, professional development, school improvement

Introduction
In a context where schools are held more and more accountable for the education they provide data-driven decision[2] making has become increasingly important. School leaders and teachers can use data to change teaching, address existing (ineffective) programs in their schools, and improve the functioning of the school in terms of increased student achievement (Young, 2006).
Data use by teachers and school leaders is often narrowly defined to the use of assessment data, particularly data from standardized assessments or national tests that is quantifiable. This view is reinforced by some policy-makers and researchers who focus on aggregated standardized test results as the primary source of data about schools, and disregard other forms of data such as the quality of instruction in classrooms. We use a broader definition of data, which encompasses the multiple types of data that teachers and school leaders need for decision-making. This includes *input* data, such as the demographics of the student population; *process* data, such as data on the quality of instruction; *outcome* data, such as student test scores (Ikemoto & Marsh, 2007). Data can thus be defined as all the information that teachers and schools need for decision-making (e.g. relevant information on students, schools, school leaders and teachers, such as assessment results and teacher surveys). A significant purpose of using data in schools is to improve student learning.

[1] Corresponding author: Dr. Kim Schildkamp, University of Twente, Faculty of Behavioral Science Department of C&O, P.O. Box 217, 7500 AE Enschede, The Netherlands, P: +31 53 489 4203, E: k.schildkamp@utwente.nl
[2] The terms data-driven decision making, data-based decision making and data use are used interchangeable.

From research in a few best practice schools we already know that teachers can use data to improve the functioning of the school in terms of increased student achievement (Lai, McNaughton, Timperley & Hsiao, 2009; Wohlstetter, Datnow & Park, 2008). However, we also know that most teachers do not use data properly, or do not use data at all (Schildkamp & Teddlie, 2008; Schildkamp & Kuiper, 2010). A majority of decisions by teachers are taken on intuition and limited observations (Ingram, Louis & Schroeder, 2004). Valuable time and resources are lost with the implementation of new curricula, which for example do not coincide with the needs of the students (Earl & Katz, 2006).
Although the attention given to the use of data in education has recently vastly increased, the field of data-based decision making is relatively new, and more research is urgently needed (Wayman & Stringfield, 2006, Kerr et al., 2006), especially into how we can support and improve data use practices within schools.
One essential aspect of using data is that teachers have time for teacher collaboration around data use (e.g. devote frequent and substantial time to reviewing data and planning accordingly) (Burbank & Kauchak, 2003; Wayman, 2005; Wohlstetter et al., 2008; Young, 2006). Schmoker (2003) states, for example, that using data should be a team effort. Based on data, a team should identify a school's strengths and weaknesses. Lessons and strategies targeted to improve student learning should be shared, produced, tested and refined. Huffman and Kalnin (2003) state that collaboration is essential for reducing the isolation of the profession, for enhancing individual teacher's professional growth, and it can have a positive impact on schools and students. It can increase teachers' knowledge on teaching, improve their teaching, and increase connections with other educators. Moreover, it can help teachers to get outside their own classroom and participate in discussion on school-wide issues.
One promising way to increase teacher involvement and collaboration as well as enhancing the effectiveness of data-driven decision making may therefore be, as stated by (Wayman, Midgley & Stringfield, 2005; 2006), setting up collaborative data teams within schools. These teams can consist of teachers and school leaders who analyze and use data to improve educational practice. Collaboration helps teachers to learn from each other how to use data, and allows for a fertile exchange of ideas and strategies (Wayman, 2005; Wohlstetter et al., 2008). This study therefore, focuses on supporting teachers in the use of data by setting up collaborative data teams (e.g. a team of teachers who collaboratively use data to solve a certain educational problem) within schools. The following research questions guided the study:

1. How do collaborative data teams function?
2. Which factors influence the work of the data team?

Theoretical framework

Teacher collaboration and data-driven decision making

Collaborative data teams are based on two key elements: the use of data and teacher collaboration. A large amount of research is available on teacher collaboration (or teacher teams or professional learning communities or teacher learning communities). Collaboration refers to a team of teachers that have a common goal, actively negotiate this goal, and try to come to a solution (Handelzalts, 2009). Collaboration is an essential

for sustainable school improvement. Teacher collaboration can have several advantages. Handelzalts (2009) summarized several of these advantages, including that it can:

- contribute to building a culture of collaboration and deliberation (e.g. lead to increased communication and interaction between teachers)
- redistribute authority (e.g. increase in decision making in different levels of the school, which gives access to new ideas and encourages a sense of ownership)
- lead to broader participation in decision making processes
- be more rewarding to the people involved, because the team structure gives teachers more emotional and moral support by extending their role, stimulates them intellectually
- lead to an increased sense of efficacy, because through interaction about teaching, teachers can improve their practice and gain a better sense of efficacy (e.g. a crucial element in improving teaching)
- lead to greater collective responsibility for teachers learning, because through interaction and collaborative work, the common goals are strengthened (e.g. causing a powerful effect on school performance).

Several of the studies reviewed by Handelzalts (2009) show that teacher collaboration can lead to school improvement in the sense that working in these teams is beneficial for student learning (see for example: Louis & Marks, 1998).
Not only teacher collaboration, but also data decision-making can lead to sustainable improvements in student learning and achievement. Several studies show that data-driven decision making (teachers making changes in the classroom based on data to improve instruction) can lead to school improvement in terms of increased student achievement levels (Campbell & Levin, 2009; Cawelti & Protheroe, 2001; Lai, McNaughton, Amituanai-Toloa, Turner, & Hsiao, 2009; Lai, McNaughton, Timperley & Hsiao, 2009; Wohlstetter, Datnow & Park, 2008). A recent synthesis of the literature on professional learning that makes a difference to student learning and achievement found that schools that used data to inquire into the effectiveness of their teaching and school practices made significant improvements in achievement (Timperley, Wilson, Barrar, & Fung, 2007).

Collaborative data teams
This study focuses on collaborative data teams. Collaborative data teams are teams of teachers (4-6) and (assistant) school leaders (1-2), who collaboratively use data to solve a certain educational problem within the school using a structured approach. The focus is on solving a problem and not on finding a problem. The process starts with questions teachers have. School leaders are part of a data team as they often have a different perspective on the educational problem to be solved, and they can bring new hypotheses to the table. Furthermore, participation of a school leader will ensure implementation of the action plan at the end of the process. The team does not have to explain anymore what they have been doing and why they want to implement certain measures. Collaborative data teams can be seen as a form of professional development with the ultimate goal of school improvement. Teachers learn in a team how to systematically use data to solve certain problems within the schools in order to improve the functioning of the school. Collaborative data teams can be seen as an effective way to improve student achievement as it combines the advantages of teacher collaboration, which studies show can lead to

school improvement (cf. Handelzalts, 2009) with the advantages of data-driven decision making, which also has the potential to lead to school improvement (Campbell & Levin, 2009; Cawelti & Protheroe, 2001; Lai, McNaughton, Amituanai-Toloa, Turner, & Hsiao, 2009).

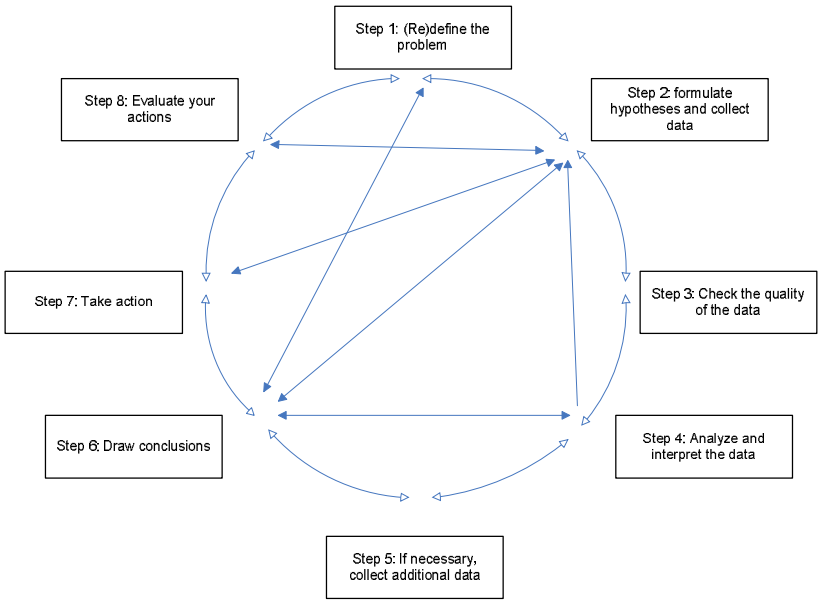

Figure 1 Collaborative data team procedure

The collaborative data teams in this study will all work according an iterative and cyclic procedure (developed based on Earl & Katz, 2006), consisting of eight steps (see also Figure 1):

1. Identifying the educational problem: the team decides on which educational problem and goals they want to focus their efforts
2. Generating hypothesis and collecting data: the team develops hypotheses and collects data to either confirm or reject the hypotheses. Several data can be collected, such as data assessment data, inspection reports, and final examination results. This step can involve simple data collection (e.g. one source of data, from one time point, from one specific group) to more complex data (e.g. more data sources, trend data, from different groups.
3. The team assesses the quality of the collected data. Are the collected data valid and reliable?
4. The team analyses and interprets the data (e.g. summarizing, calculating, comparing). This can involve simple data analyses (e.g. descriptive analyses) as well as more sophisticated analyses (e.g. regression and multilevel analyses).
5. The team decides if they know what causes the problem or that they need to collect additional data, for example by conducting classroom observations, or that they need to refine their problem statement, or generate new hypotheses.
6. The team draws conclusions based on all the collected data
7. The team takes action based on the data. The team describes the actions, and the goals that go with the actions. The team makes team members responsible for

implementing the actions, and determines which resources are available for implementing the actions. The team also thinks of ways to monitor the implementation of the actions, sets deadlines, and determines which data is needed to establish the effectiveness of the implemented actions.

8. The team evaluates the actions implemented. Are the actions effective? Are the goals met? Are the problems solved and is the team satisfied? To evaluate the actions, new data needs to be collected (step 2). This process continues until the priorities are met and the goals are reached. Then the team can go back to step 1 to work on a new problem and goals.

Next to these activities, the functioning of data teams can also be characterized by the relevance of the discussions the team has (based on Henry 2010; Ikemoto & Marsh, 2007). Henry (2010) used instructional relevance (based on Cohen & Ball, 1999; 2001) to characterize and analyze team conversations of data teams. Based on her study, Henry (2010) concluded that this is a valuable concepts in analyzing and understanding teachers' collaborative work in data teams and how it may support efforts to improve teaching and learning in schools. In order to come to a nuanced view of the work of data teams it is essential to understand what (relevance) is being discussed. Moreover, patterns of differences across teams may provide information on why some data teams are able to solve the educational problem they are working on, whereas others are not.

Instructionally relevant conversations are conversations about relationships between teachers, students and content. Conversations that lack relevance are, for example, about paperwork, scheduling or personal stories (Henry 2010). Suppovitz (2002) studied the effectiveness of several teams and found that instructional relevant conversation were more prevalent in the effective teams, whereas the less successful teams spend more time on administrative and paper work. The literature review conducted by Henry (2010) suggests that teams that are achieving student learning gains are teams where instructional relevant conversations are common. The less effective teams have less instructional relevant conversation.
Relevance in this study relates to whether or not the focus of the team conversations is on the educational problem to be solved. Conversations with no relevance do not focus on the problem to be solved. It may focus on personal storytelling, other professional responsibilities, or agenda setting. Conversations with low relevance focus on the problem, but without consideration of how the problem is influenced by school and classroom mechanisms. The focus is on possible causes of the problem outside the school (e.g. low entry level of students, bad performing primary schools). Conversations that focus on the problem, but without consideration of how the problem is influenced by classroom mechanisms, are characterized as minor relevant conversations. The focus is on possible causes of the problem at school level (e.g. transfer rules from one grade to another are not strict enough, too little hours available to teach the subject effectively). Moderate relevant conversations focus on the problem and how it is being influenced by classroom mechanisms, but without consideration of how one's own functioning relates to the problem. The problem is for example caused by students that are lazy, other teachers that do not spend enough time on the topic, or by the fact that students are not interested in the subject that they teach. Highly relevant conversations focus on how the

problem is related to one's own functioning, and the interaction between students, teachers, and the content taught (e.g. which instructional strategies can I use to help my students understand mathematics better). These conversations are identified as highly relevant, as research show that teachers have a central role in realizing sustainable, significant and coherent improvement of teaching and learning practices (c.f. Fullan, 2001; Hopkins, 2001; Levin, 2008), and they are the ones who have to change their practice in order to improve student achievement.

It is hypothesized that collaborative data teams, which have more conversations that can be described as relevant will be more effective in solving the educational problem and improving education as teams with less relevant conversations. Relevant inquiry focused process can lead to improved student achievement (Feldman & Tung, 2001).

Factors influencing data-driven decision making

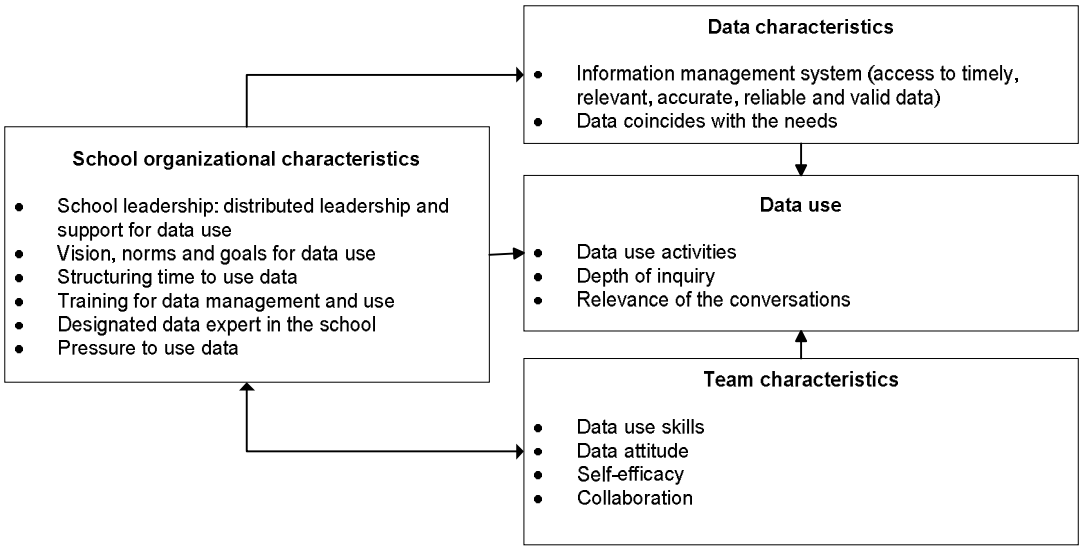

Figure 2 Factors hypothesized to influence data use

Several factors may promote or hinder the work of collaborative data teams. These factors have been identified by conducting an extensive literature review on data use, teacher collaboration, and related fields (see for example: the special issue of the American Journal of Education, 2006, 112(4); Ikemoto & Marsh, 2007; Kerr et al., 2006; Visscher, 2002; Wayman, 2005; Wohlstetter et al., 2008, Schildkamp, 2007; Schildkamp & Kuiper, 2010; Schildkamp & Teddlie, 2008, Schildkamp, et al., 2009; Vanhoof et al., 2009; Verhaeghe et al., 2009). The following factors are hypothesized to influence data use: data characteristics, respondent and team characteristics, and school organizational characteristics (see Figure 2) (for an extended description of all variables, see Schildkamp & Kuiper, 2010).

The Dutch context

Dutch schools traditionally have considerable autonomy. They have always been free to choose the religious, ideological and pedagogical principles on which they base their education, as well as how they choose to organize their teaching activities (Ministerie van

Onderwijs, Cultuur & Wetenschappen, 2000). The Netherlands do have an inspectorate, which holds schools accountable for their education.

On the one hand there are explicit expectations of schools using data to monitor and improve their quality, as school inspections are based on quality assurance and school self-evaluation data. On the other hand, the use of, for example school self-evaluation data, is usually perceived by school staff as a task of school management. Teachers feel that their primary task is teaching, and this does not include the use of (self-evaluation) data (Schildkamp, 2007). Moreover, since it is difficult to sanction schools, school staff may feel less pressured to use data to improve. This context may influence which data the data teams can use. In the Netherlands, several data sources are available within schools, including:

- School inspection data: the inspection generally judges the school on the educational processes based on the supervision framework discussed above. Schools are also judged on the basis of their output, including the percentage of grade repeaters in lower and higher levels of secondary education, as well as the average grade for the final exams of each track.
- School self-evaluation data: school self-evaluation can be defined as a procedure involving systematic information gathering initiated by the school with the intention to assess the functioning of the school for supporting decision-making, organizational learning, and for fostering school improvement (Schildkamp, 2007). In the Netherlands, more than 70 different instruments for school self-evaluation are available (The Standing International Conference of Central and General Inspectorates of Education, 2003). Schools use different types of school self-evaluation instruments, but most school self-evaluations include management and teacher questionnaires on the functioning of the school. Sometimes, school self-evaluation instruments also include student and parent questionnaires.
- Data on intake, transfer and school leavers: schools keep records on the intake of students, transfer (e.g. grade repeaters), and school leavers (with or without a diploma).
- Final examination results: at the end of secondary education, students have to pass a final examination. Dutch upper secondary education (which is the context of this study) encompasses two school types: HAVO (general secondary education) and VWO (university preparatory education). Both are for pupils from the age of 12 years and up. VWO diplomas qualify pupils to enter university or higher professional education. The HAVO diploma provides entry to higher professional education, but also to the fifth year of VWO or to upper secondary vocational education. The curriculum is organized around four subject profiles (nature and technology, nature and health, economics and society and culture and society). Pupils complete the requirements for at least one profile. This involves instruction in an average of 15 subjects, some of which are examined internally by the school, some of which are examined both internally and during a final national exam. Final examination results thus include the results of internal and external assessments.
- Assessment results: students in the Netherlands are assessed on a regular basis. Assessments include both oral and written assessments, and are usually administered by teachers.
- Student questionnaire data and focus groups: some schools administer student questionnaires, to gather information on the students' perception on the functioning of

teachers, school leaders and the school in general. Some schools also use student focus groups for these purposes.

- Parent questionnaire data and focus groups: some schools administer parent questionnaires, to determine the needs of their stakeholders, to gather information on the parents perception on the functioning the school. Some schools also use parent focus groups for these purposes.

Method

To explore the functioning of collaborative data teams a pilot is being conducted. The functioning of collaborative data teams is being studied in five Dutch schools for upper secondary education. These schools approached the university themselves. They had heard that the university wanted to pilot a new form of professional development, collaborative data teams, and they would like to participate. Three data teams were established in these schools. The results described in this paper are the results of the first year of the two-year pilot. Although data from only three teams, which volunteered to participate, do not permit extensive generalizations of the findings, they may offer valuable new insights into potential of collaborative data teams to stimulate data use in Dutch schools. Based on these results, future research can focus, for example, on large scale implementation of collaborative data teams, and its effects.

Respondents

Table 1 *Respondents*

Team	Participants	Educational problem
Data team 1 (Carmel)	• 2 mathematics teachers (school 1) • 2 mathematic teachers (school 2) • 1 mathematic teacher (school 3) • 1 data expert (school 1-3) • 1 assistant school leaders (school 1)	Low mathematics achievement in the lower grades of secondary education
Data team 2 (Brug)	• 6 mathematics teachers (school 4) • 1 data expert (school 4 and 5) • 1 assistant school leader (school 4)	Low mathematics achievement in the third grade of secondary education
Data team 3 (Waal)	• 2 English teachers and mentor (school 5) • 1 Dutch teacher and coordinator Havo (school 5) • 1 History teacher and mentor (school 5) • 1 Music and arts teacher and mentor (school 5) • 1 data expert (school 4 and 5) • 1 assistant school leader (school 5)	High percentage of grade repeaters in the third grade of secondary education

For this study, three collaborative data teams from five schools (see Table 1) were followed during one school year. The problems these teams tried to solve were:

- Low mathematic achievement in the lower grades of secondary education (two teams)
- High percentage of grade repeaters in the third grade of secondary education

All teams were guided by a researcher. The role of the researcher was to facilitate the work of the data teams, for example in analyzing the data.

Data collection and instruments

Observations (see Table 2) were used to study the work of the collaborative data teams. Observations make it possible to study actual behavior of individuals and groups (Baarda & De Goede, 1997), and to gain insights in the communication and interaction within a group (Wester, 1987).

Open observations were used, which revolved around giving a detailed description of the functioning of the data teams, and the influence of the team characteristics, school organizational characteristics and data characteristics. All team meetings were followed and audio recorded and processed on the basis of the themes discussed in our theoretical framework. The observations were converted into transcripts and analyzed by using the qualitative data analyses program Atlas/ti, which allows for coding all observation fragments, relating the coded fragments to each other, and comparing the codes of teams and respondents. The observation data were coded based on the theoretical framework developed. The coded data were then used to develop detailed report cases for each school according to a common outline. This facilitated a cross site analysis.

Also, all participants of the data teams were interviewed (see Table 2). This interview schedule was developed based on the theoretical framework and interview schedules developed by Cousins, Ross and Maynes (1994), and Handelzalts (2009) to study data use and teacher collaboration. Questions were formulated with regard to: the way participants experienced the data teams and using data in these teams, and which they though stimulated or hindered their work (e.g. the influencing factors). These interviews took place after 3 or 4 meetings.

Interviews were converted into transcripts and analyzed by using the program Atlas/ti. The interview data were also coded based on the theoretical framework developed. The coded data were then added to the developed detailed report cases for each school.

Table 2 *Number of respondents interviewed and meetings observed*

		Interviews (assistant) school leaders	Interviews teachers	Interviews quality care manager/ data expert	Data team meetings observed
Team (Carmel)	1	1	5	1	8
Team (Brug)	2	1	6	1	5
Team (Waal)	3	1	5	1	5

Reliability and validity

In this study reliability was fostered by using a systematized approach to data collection that is consistent with the research questions. We used a protocol, which described the research questions, data collection method and instruments, and analysis planned. Reliability was further enhanced by audio taping all data team meetings and the interviews conducted, which allowed for thorough analyses of the data. Also, all the instruments used were based on our theoretical framework and existing instruments. Finally, a second researcher coded a part of the observation data and interview data. We found Cohen's Kappa to be 0.75, which is more than sufficient (Israel, 2008).

Internal validity was enhanced by highlighting major patterns of similarities and differences betweens respondents' experiences and beliefs in one table. For enhancing construct validity, multiple sources of evidence or triangulation (i.e., interviews and observations) (see also Table 2) were used. External validity was realized by providing case-specific and cross-case thick descriptions (also including citations of respondents), and describing the congruence with the theoretical framework.

Results

Activity level

Team 1 focused on the educational problem of low mathematic achievement in the first and second grade of secondary education (students aged 12-14), which they all considered to be a problem. The team decided to narrow the problem to a lack of skills in completing narrative mathematic assignments (e.g. Alice is walking home from school. First she drops by her friend's house, which is located 0.5 mile from her school. Next, she walks to the grocery store, which is located 0.3 mile from her friend's house. Then she walks the final 0.2 mile home. How many miles did Alice walk?).

The team collected assessment data to further analyze the problem. They discovered that the problem was not as big as they thought is was. After a few lessons, all students are able to complete these types of assignments. They also discovered a new problem in the assessment data: students were not able to complete percentage and fracture assignments. They decide to explore this educational problem further.

The next couple of meetings show an iterative process of generating hypothesis, collecting and analyzing data, and rejecting and accepting hypothesis. Several hypotheses of the team were rejected based on data (e.g. data show that no relation could be found between low achievement results and the mathematic curriculum of certain primary schools, which teachers taught did not relate well to the curriculum of their school; a previously administered motivation questionnaire showed that no relationship could be found between motivation and mathematic achievement). All of the first hypotheses of the team were rejected.

However, instead of having a de-motivating effect on the team, all the team members indicated that they found this very enlightening. One of the teachers stated, for example, "For 10 years now I have blamed certain primary schools for our problems, but it looks like I don't have to worry about these primary schools anymore".

After a few meetings of rejecting hypotheses based on data and generating new hypotheses, the team found out what caused the educational problem. They were able to accept two important hypotheses. The first hypothesis related to the entry level of students. Assessment data (Cito) from the primary schools that gives information on the entry level of the students showed that several students had very low mathematic skills. Another hypothesis confirmed, based on two assessments administrated in the beginning of the school year and again after a couple of months, that students were able to complete the assessment successfully right after they had been taught how to solve percentage and fracture assignments, but they forgot these skills a couple of months later.

Based on these hypotheses the team came up with two measures to solve this problem: (1) the data show that students need repetition when it comes to learning how to solve percentage and fracture assignments. However, mathematic teachers also have to teach

other topics such as geometry and can not spend entire lessons on repetition of percentage and fracture solving skills. After a brainstorm session, they came up with the following solution: they decided to start each mathematic lesson with a 5 minute pop quiz in percentages and fractures. This would only take a limited amount of their time and they would still be able to provide students with the repetition they need. Moreover, they would include a percentage and/or fracture assignment in each assessment throughout the whole school year. (2) For the low entry level students the team found several online programs students can use to practice. The teachers decided to advice low performing students (which they were able to identify early based on entry level scores) to practice after lessons and at home. The next data team meeting was scheduled at the beginning of the next school year, when the teachers would further discuss the implementation of their measures as well as how to evaluate the effectiveness of their measures.

Team 2 also focused on the low mathematics achievement results, but they focused on the third grade (age 14/15) results. However, at the first meeting it soon became clear that the school leader of this school considered this to be a problem, but that the teachers wanted to use the data team to show the school leader that it is not a problem.
The first couple of meetings of this team all concerned the discussion on what type of data the team needed to determine whether their mathematic achievement results were too low or not, and what the criterion should be for identifying it as a problem or not. After a lot of discussion (the school leader wanted to set the threshold at 80% and the teachers thought 60% to be realistic), they decided that at least 70% of the students should pass mathematics on their report card (although some of the teachers stated that they would still be happy as 60% would pass, because mathematics is a difficult subject, but they were still willing to work on improvement if this would be the case).
It was not until meeting 5 that the entire team finally agreed that the mathematic results were too low (passing rates varied between classes from 46% to 75%, but several classes did not meet the threshold of 70%).
The team generated several hypotheses, including that the entry level of several students is too low; students receive too little mathematic lessons due to a high degree of student as well as teacher absence; and achievement declined with the implementation of the 60 minutes roster. For each of the hypothesis it was determined which data the team needed and where to find these data. One of the teachers promised to collect these data and to bring it to the next meeting in the beginning of the new school year.

Team 3 focused on an entirely different problem: the large number of grade repetition in the third grade. At the first meeting, it became clear that the whole team considered this to be a big problem. The next couple of meetings focused on generating hypotheses, collecting data and rejecting or accepting hypotheses. The problem soon proofed to be very complex. One of the hypotheses of the team was that the transfer rules of the school were not strict enough. Students with four or five insufficient marks on their report card were still allowed to transfer from the second to the third grade. These students were not able to compensate for their disadvantage and had to repeat the third grade (instead of the second grade). Data from the report cards showed that this was true, but only for 6 students. Another hypothesis was that several students ignored their advice to continue secondary education at a lower level, and ran into problems in the third grade. The team

had to reject this hypothesis. The students with the advice to continue at a lower level were doing fine in the third grade. Based on student motivational questionnaire results the hypothesis that students were not motivated enough, was rejected as well. The questionnaire results did show that student needed help in planning and needed more feedback from teachers on their planning activities. At the next meeting, at the beginning of the new school year, the team needs to come up with new hypotheses.

Relevance of conversations
The data team conversations of team 1 started out with low relevance. The problem of low mathematic achievement in the lower grades of secondary education was, for example, caused by low performing primary education. However, when these hypotheses were all rejected based on data, the team switched from having low relevance conversations to having highly relevant conversations. This team had the highest number of highly relevant conversations. The team focused on "changes they could make in their classrooms to improve mathematic achievement". When data showed that children need repetition when it comes to percentage en fracture assignments, they implemented the pop quizzes and the use of the online program.

Team 2 started out with several conversations with no relevance to the problem to be solved (low mathematic achievement in the third grade). There appeared to be some problems between the school leader and teachers and the first couple of meetings were mainly focused on these problems and on arguments whether or not the problem was indeed a problem (the teachers did not think so). After a couple of meetings the conversations switched to low relevance conversations. According to the teachers, the problem is caused by the facts that "mathematics is just a very difficult subject" and "the students we have at this school are just not capable enough". This team is still focusing on these conversations with low relevance. The team is not any closer to solving the problem on low mathematic achievement.

Team 3 started out with minor relevance conversations. The problem of grade repetition was caused by the policy of the school with regard to grade transfer. Contrary to school 1, this data team came up with a lot of school level hypotheses (e.g. minor relevance), before switching to highly relevant conversations. It was not until the results of a school level questionnaire on student motivation showed that only the classroom level items influenced grade repetition, that this data team saw that they needed to focus on classroom level explanations and that they needed to look at their own functioning as well.

Supporting and hindering factors
In terms of the data characteristics, all three teams had access to the data they needed. The quality care manager (e.g. the data expert) played a crucial role here, as she had access to the school level information management systems as well as all kinds of other data.

Of the school organizational characteristics, school leadership seems to be essential. Teams 1 and 3 were facilitated and supported in their work by the school leader. As one

of the teachers stated "Sometimes our schedule is cleared (e.g. cancellation of lessons) so that we can participate in the data use meeting. This is special. Our school really tries not to cancel any lessons. If lessons are cancelled for this, this means that it is very important." On the contrary, in data team 2, the relationship between the school leader and teachers in team 2 hindered the work of the data team.

Also, having a clear goal is essential. Team 1 (improving mathematic achievement) and team 3 (reducing grade repetition) had clear goals, and everyone in the team agreed that these were important goals. According to one of the teachers, who expresses the importance of solving the problem as well as using data: "I wanted to participate because I experience the problem myself. I like the fact that we are using data. Data are objective and do not include the posturing we experience a lot in education". The goals for team 2 were less clear, as the team did not really agree on the problem. This team did not consider it to be a problem that nearly half of their students failed mathematics. Probably because team 2 did not really feel the importance of "their problem", it was difficult to schedule time for meetings with this team. In all the data teams, structuring time for data use was an issue, but in teams 1 and 3 all team members tried really hard to make time for the meetings.

Data use skills was a team characteristics almost all team members in all three teams lacked. The teachers and school leaders were not ably to analyze the collected data. This had to be done by the quality care manager and researcher. The role of the researcher is also essential, as expressed by one of the teachers: "she guides the process and gives us tips on the different steps, and she analyzes data, which we can't."

What is striking is that all team members had a positive attitude towards data use. Even the team members of team 2 expressed that in education you can not "just rely on your guts and instincts; you have to look at hard data as well". However, although all team members expressed a positive attitude towards the use of data, this team was not functioning very well. Collaboration between team members may have hindered their work. Because of problems between the school leader and the teachers, the work of the data team focused more on this relationship than on the actual problem. In the other teams, the team members worked together more effectively. As on of the teacher stated: "People work together, appointments are kept, everyone is actively involved, people listen to each other and are not afraid to speak up. We all work together at an equal level."

Conclusion and discussion

The results of this study show that teams 1 and 3 are clearly more effective in their work (e.g. solving the problem). Team 1 had the highest number of highly relevance conversations, and was also closest to actually solving the problem. The results of this study also point to some important issues we need to address if we want to support data-driven decision making in schools. None of the team members possessed the skills to employ data analyses on the collected data. They were not familiar with programs such as SPSS, and a majority did not even now how to use Excel for analyses. Therefore, the researcher had to assist in employing the analyses (e.g. regression analyses to determine the relationship between certain factors). Research has revealed that often even simple

conceptions and representations are interpreted incorrectly within schools (Earl & Fullan, 2003). A sufficient level of assessment literacy is a prerequisite for correct understanding. If not, proper support initiatives should be foreseen (e.g. use a facilitator who can do the analyses as in this study or educate school staff in more complex data analyses).

Also, a supporting and facilitating school leader seems to be essential for the work of data teams. Teams 1 and 3 were more effective, and these teams had a supportive and enthusiastic school leaders, who stressed the importance of data use. These school leaders also expressed a clear vision for data use. Other research also shows, that teachers are more willing to participate in decision making if they perceive their relationship with the school leader as more open, collaborative, facilitative, and supportive (Smylie, 1992).

Also, starting with a problem and clear goals which all the data team members agree upon is essential. Data use is more likely to be successful if a team starts with an agreed upon problem and does not have to go through mounts of data to find a problem. Other research (Datnow et al., 2007; Earl & Katz, 2006; Feldman & Tung, 2001; Kerr et al., 2006; King, 2002; Sharkey & Murnane, 2006; Wayman & Stringfield, 2006b; Wayman et al., 2007; Wohlstetter et al., 2008; Young, 2006) also stresses the importance of having a clear goal for data use.

The results of this study further show that the work of data teams is essential in improving education. All the teachers in the three teams started out with hypotheses based on their beliefs, which were proven to be incorrect. As about 40% of teachers' decisions are based on their beliefs (Ingram et al., 2004), a lot can be gained here by the use of data.

Although the (tentative) results described here are based on a small pilot, we do belief that the data team approach is an effective approach in supporting schools in the use of data to improve education. The data team approach seems to be effective in two different manners (1) actually solving certain problems within the school and thereby improving education and (2) educating teachers and school leaders in how to use data to improve education (e.g. a form of professional development). Therefore, the next step of this study (after completing the pilot) is to expand the data team procedure in more schools, primary as well as secondary schools.

References

Baarda, D. B., Goede, M. P. M. de., & Teunissen, J. (1997). *Basisboek kwalitatief onderzoek [Handbook for qualitative research]*. Houten: Educatieve Partners Nederland BV.

Burbank, M.D., & Kauchak, S (2003). An alternative model for professional development: Investigations into effective collaboration. *Teaching and Teacher Education*, 499-514.

Campbell, C., & Levin, B. (2009). Using data to support educational improvement *Educational Assessment, Evaluation and Accountability, 21*, 47-65.

Cawelti, G., & Protheroe, N. (2001). *High student achievement: How six school districts changed into high-performance systems.* Arlington, VA: Educational Research Service.

Cousins, J.B., & Ross, J.A., & Maynes, F.J. (1994). The reported nature and consequences of teachers' joint work in three examplary schools. *The Elementary School Journal,* 94*(4),* 441-465.

Datnow, A., Park, V., & Wohlstetter, P. (2007). *Achieving with data. How-high performing school systems use data to improve instruction for elementary students*. San Francisco: Center on Educational Governance University of California.

Earl, L. Fullan, M. Using Data in Leadership for Learning. Cambridge Journal of Education, 33 (3). 383-394, 2003.

Earl, L.M., & Katz, S. (2006). *Leading schools in a data-rich world. Harnessing data for school improvement.* Thousand Oaks: Corwin press.

Feldman, J., & Tung, R. (2001). *Whole school reform: how schools use the data-based inquiry and decision making process.* Paper presented at the American Educational Research Association conference, Seattle, April 10-14.

Fullan, M. (2001). *The new meaning of educational change*. New York: Teachers College Press.

Handelzalts, A. (2009). *Collaborative curriculum development in teacher design teams.* Enschede: Universiteit Twente.

Henry, S. F. (2010). *Talk that talk: A qualitative exploration of relevance and depth in elementary teachers' team conversations*. Paper presented at the American Educational Research Association Conference. April 29 – May 4, 2010.

Hopkins, D. (2001). *School improvement for real: Educationa; improvement for real.* London: Routledge Falmer.

Huffman, D., & Kalnin, J. (2003). Collaborative inquiry to make data-based decisions in schools. *Teaching and Teacher Education, 19*(6), 569-580.

Ikemoto, G. S., & Marsh, J. A. (2007). Cutting through the "data driven" mantra: Different conceptions of data-driven decision making. In P.A. Moss (Ed.), *Evidence and decision making* (National Society for the Study of Education Yearbook, Vol. 106, Issue 1, pp.105-131). Chicago: National Society for the Study of Education.

Ingram, D., Louis, K. S., & Schroeder, R. (2004). Accountability policies and teacher decision making: barriers to the use of data to improve practice. *Teachers College Record, 106*(6), 1258-1287.

Israel, D. (2008). Data analysis in business research: A step by step nonparametric approach. London: Sage Publications.

Kerr, K. A., Marsh, J. A., Ikemoto, G. S., Darilek, H., & Barney, H. (2006). Strategies to Promote data Use for Instructional Improvements: Actions, Outcomes, and Lessons from Three Urban Districts. *American Journal of Education, 112*, 496-520.

King, M. B. (2002). Professional development to promote schoolwide inquiry. *Teaching and Teacher Education, 18*(3), 243-257.

Lai, M. K., McNaughton, S., Amituanai-Toloa, M., Turner, R., & Hsiao, S. (2009). Sustained acceleration of achievement in reading comprehension: The New Zealand experience. *Reading Research Quarterly, 44*(1), 30-56.

Lai, M.K., McNaughton, S., Timperley, H., & Hsiao, S. (2009). Sustaining continued acceleration in reading comprehension achievement following an intervention. *Educational Assessment, Evaluation and Accountability, 21*(1), 81-100.

Levin, B. (2008). *How to change 500 schools. A practical and positive approach for leading change at every level.* Cambridge: Harvard Education Press.

Louis, K. S., & Marks, H. M. (1998). Does professional community affect the classroom? Teachers' work and student experiences in restructuring schools. *American Journal of Education, 106*(4), 532-575.

Ministerie van Onderwijs, Cultuur & Wetenschappen (2000). *Wet op het onderwijstoezicht [Education supervision act]*. Den Haag: SDU.

Schildkamp, K. (2007). The utilisation of a self-evaluation instrument for primary education. Enschede: Universiteit Twente.

Schildkamp, K., & Kuiper, W. (2010). Data Informed Curriculum Reform: Which data, what purposes, and promoting and hindering factors. *Teaching and Teacher Education, 26*, 482-496.

Schildkamp, K., & Teddlie, C. (2008). School performance feedback systems in the USA and in the Netherlands: a comparison. *Educational Research and Evaluation,14*(3), 255–282Schildkamp, K., Visscher, A., & Luyten, H. (2009). The effects of the use of a school self-evaluation instrument. *School Effectiveness and School Improvement, 20*(1), 69-88.

Schmoker, M. (2003). First things first: demystifying data analysis. Educational Leadership, 60(5), 22–24.

Sharkey, N. S., & Murnane, R. J. (2006). Tough choices in designing a formative assessment system. American Journal of Education, 112, 572–588.

Smylie, M.A. (1992) Teacher Participation in School Decision Making: Assessing Willingness to Participate. Educational Evaluation and Policy Analysis. 14 (1), 53-67 1992.

Supovitz, J. A. (2002). Developing communities of instructional practice. *Teachers College Record,104*(8), 1591-1626.

Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). *Best evidence synthesis iterations (BES) on professional learning and development.* Wellington: Ministry of Education.

Vanhoof, J., Van Petegem, P., & De Maeyer, S. (2009). Attitudes towards school self-evaluation. Studies in Educational Evaluation, 35, 21-28.

Verhaeghe, G., Vanhoof, J., Valcke, M., & Van Petegem, P. (2009). Using School Performance Feedback: Perceptions of Primary School Principals. School Effectiveness and School Improvement. (in press).

Visscher, A. J. (2002). A framework for studying school performance feedback systems. In A. J. Visscher, & R. Coe (Eds.), School improvement through performance feedback (pp. 41–72). Lisse: Swets & Zeitlinger B.V.

Wayman, J. C. (2005). Involving teachers in data-driven decision making: using computer data systems to support teacher inquiry and reflection. Journal of Education for Students Placed at Risk, 10(3), 295–308.

Wayman, J.C., & Stringfield, S. (2006). Data Use for School Improvement: School Practices and Research Perspectives. *American Journal of Education, 112*, 463-468.

Wayman, J. C., Midgley, S., Stringfield, S. (April 11–15, 2005). Collaborative teams to support data-based decision making and instructional improvement. Paper presented at the American educational research association conference, Montreal.

Wayman, J. C., Midgley, S., & Stringfield, S. (2006). Leadership for data-based decision making: collaborative educator teams. In A. Danzig, K. Borman, B. Jones, & B. Wright (Eds.), New models of professional development for learner centered leadership (pp. 189–205). Erlbaum.

Wayman, J. C., Cho, V., & Johnston, M. T. (2007). The data-informed district: A district-wide evaluation of data use in the Natrona County School District. Austin: The University of Texas.

Wester, F. (1987). Strategieën voor kwalitatief onderzoek. Muiderberg: Coutinho.

Wohlstetter, P., Datnow, A., & Park, V. (2008). Creating a system for data-driven decision-making: Applying the principal-agent framework. *School Effectiveness and School Improvement, 19*(3), 239-259.

Young, V.M. Teachers' Use of Data: Loose Coupling, Agenda Setting, and Team Norms. American Journal of Education. 112. 2006 521-548.