Teaching and Teacher Education 67 (2017) 509–521

Contents lists available at ScienceDirect

Teaching and Teacher Education

journal homepage: www.elsevier.com/locate/tate

The quest for sustained data use: Developing organizational routines

Mireille D. Hubers*, Kim Schildkamp, Cindy L. Poortman, Jules M. Pieters

University of Twente, PO Box 217, 7500 AE, Enschede, The Netherlands

HIGHLIGHTS

- Schools' organizational routines for data use were scarce and developed little over time.
- Especially the ostensive aspect (e.g., policy and vision for data use) was underdeveloped.
- Interventions for data use should target the development of organizational routines.

ARTICLE INFO

Article history:
Received 8 January 2017
Received in revised form
25 April 2017
Accepted 6 July 2017
Available online 7 August 2017

Keywords:
Data use
Professional development
Organizational routines
Sustainability

ABSTRACT

The data team intervention was designed to support schools' data use. The sustainability of schools' data use was investigated by studying the schools' development of the ostensive and performative aspects of organizational routines for: engaging in the data team intervention, acting upon their data team's improvement plan, and using data for school development and instruction. Six Dutch secondary schools participated in this longitudinal mixed-methods study. Data were collected through questionnaires, policy documents, and interviews. Results indicated that schools struggled to develop organizational routines for data use, especially the ostensive aspects. This illuminated the process by which schools did not sustain their use of data. The findings showed that interventions for data use should more clearly target the development of organizational routines.

1. Introduction

Data-based decision making in education has been emphasized globally in recent years (Datnow & Hubbard, 2015; Datnow, Park, & Kennedy-Lewis, 2013). Data-based decision making, or data use for short, is important because decisions informed by data are more likely to be effective than decisions based on intuition and experience (Schildkamp & Poortman, 2015; Schildkamp, Poortman, & Handelzalts, 2015). Data can support teachers' processes of reflection and provide insight into their strengths and weaknesses. As a result, teachers may change their behavior, such as by trying out different instructional strategies, which can improve their own performance (Schildkamp & Kuiper, 2010). Overall, the use of data can improve the quality of the education provided by individual teachers, schools, or districts, which can lead to improved student achievement (Campbell & Fullan, 2006; Carlson, Borman, & Robinson, 2011; Lai & Schildkamp, 2013; Van Geel, Keuning, Visscher, & Fox, 2016).

Even though data use is associated with various benefits, previous studies have shown that most teachers do not use data to its best effect or do not use data at all (e.g., Means, Chen, DeBarger, & Padilla, 2011; Oláh, Lawrence, & Riggan, 2010). To support schools in their use of data, several professional development programs have been developed that target (a combination of) student learning, teacher learning, and organizational change (e.g., Coburn & Turner, 2011; Ikemoto & Honig, 2010; Karr, Marsh, Ikemoto, Darilek, & Barney, 2006; Marsh, McCombs, & Martorell, 2009; Schildkamp et al., 2015).

One such professional development program is the data team intervention (Schildkamp et al., 2015). Teams that work according to this intervention consist of six to eight teachers and school leaders, who learn how to use data to analyze and address an educational problem in their school (e.g., high grade retention rates). Previous research has illustrated that working with this intervention increased data team members' knowledge and skills regarding data use (Ebbeler, Poortman, Schildkamp, & Pieters, 2016), and that working with this intervention can lead to increased student achievement (Poortman & Schildkamp, 2016).

* Corresponding author.
E-mail addresses: m.d.hubers@utwente.nl (M.D. Hubers), k.schildkamp@utwente.nl (K. Schildkamp), c.l.poortman@utwente.nl (C.L. Poortman), j.m.pieters@utwente.nl (J.M. Pieters).

http://dx.doi.org/10.1016/j.tate.2017.07.007
0742-051X/

However, it is not yet known whether working with this intervention can result in lasting school-wide changes in policy, the organization of work, and work practices themselves (e.g., Coburn & Turner, 2011; Honig, 2008; Sherer & Spillane, 2011), an issue referred to as sustainability (Fullan, 2007). The extent to which data use is sustained can be studied through the development of schools' organizational routines. These routines are recurring actions that structure everyday practice in schools by supporting and focusing interactions among school staff (March & Simon, 1958; Nelson & Winter 1982). They include both the broader organizational system of the school and actual work practices (Sherer & Spillane, 2011).

Even though the development of organizational routines is key in understanding whether and how data use programs lead to sustainable changes in educational practice (Spillane, 2012), few studies have addressed the development of such routines (e.g., Coburn & Turner, 2011; Little, 2012; Spillane, 2012). Insights into schools' organizational routines is crucial for coming to a deeper understanding of the dynamic between an intervention, and the resulting on-the-ground responses and actions, such as how data are being used (Coburn & Turner, 2011; Marsh, 2012). Furthermore, these insights are critical for making an informed decision on whether a program is worth the investment of efforts and resources (Coburn & Turner, 2011). The present study aimed to investigate whether schools who implemented the data team intervention sustained their data use. Doing so involved studying how their organizational routines related to data use developed over time. This provides insight into the process through which schools did or did not sustain the use of data in their educational practice.

2. Theoretical framework

2.1. Teachers' professional development

The use of professional learning communities is seen as an important way to support teachers in rethinking their own practice and improving their teaching (Vescio, Ross, & Adams, 2008). Despite promising research results, professional development and school improvement do not automatically take place as a result of working in these communities. Previous research illustrated that the effects may be small and results are mixed (Chapman & Muijs, 2014; Lomos, Hofman, & Bosker, 2011). Therefore, one of the biggest challenges is to make sustainable school-wide changes in policy and practice (e.g., Harris & Jones, 2010; Van Veen et al., 2010).

To study this sustainability challenge, roughly two perspectives can be distinguished: the determinants and the dimensions of change (Crossan & Apaydin, 2010). The former refers to the way in which factors influence the implementation of change. Examples include the role of school climate, school leaders, teachers' beliefs, sensemaking, agency, and available time/money (Coburn & Talbert, 2006; Datnow, Park, & Wohlstetter, 2007; Desimone, 2002; Kurland, Peretz, & Hertz-Lazarowitz, 2010; Penuel, Fishman, Cheng, & Sabelli, 2011; Schildkamp & Kuiper, 2010; Visscher, 2002). The latter, the dimensions of change, refers to the process (the 'how') of change and the outcome of the change (the 'what') (Crossan & Apaydin, 2010). The process view is especially underdeveloped in the literature (Crossan & Apaydin, 2010). As this is a crucial aspect of understanding how change develops in schools, the present study used a process view. The development of schools' organizational routines was studied to determine whether sustainable changes were made while some teachers worked in professional learning communities (data teams).

2.2. Organizational routines

Much of the work in schools takes place in and through organizational routines (Nelson & Winter 1982). An organizational routine can be defined as 'a repetitive, recognizable pattern of interdependent actions, involving multiple actors' (Feldman & Pentland, 2003, pp. 96). An example of this would be a group of mathematics teachers who determine at the end of each year how their students have performed and use this information to determine how the coherence in the curriculum can be improved.

There are several reasons why the data team intervention, and other professional development programs for that matter, can alter the organizational routines around data use in a school (Coburn & Turner, 2011). First, an intervention can bring educators together in new and different combinations, which can influence the dynamics through which they both interpret the data and design corresponding actions for improvement (Coburn & Russell, 2008). For example, Marsh, Bertrand, and Huguet (2015) found that teachers' participation in data use professional learning communities played an important role in teachers' responses to data, and, subsequently, that they used those data to alter their instructional delivery. Second, an intervention can shape individual and collective beliefs; for example, ample time can be provided to openly discuss student achievement and reexamine requirements (McDougall, Saunders, & Goldenberg, 2007). Third, it can shape what educators notice and attend to by focusing their attention on certain data and not others (e.g., Ikemoto & Honig, 2010; Sherer & Spillane, 2011). Finally, those engaging in the intervention can broker their (newly gained) knowledge about data use and the corresponding actions for improvement to their colleagues (Hubers et al., 2017). This can facilitate the school staff's participation in discussions on school-wide issues and increase communication about data use and the issues those data indicate to be important (Huffman & Kalnin, 2003; Lachat & Smith, 2005).

Organizational routines illuminate the extent to which data are used because they focus our attention on the interactions among school staff (Spillane, 2012). In these interactions, teachers and school leaders negotiate what data are worth noticing and what these data mean for current practice at the school and classroom levels. This clarifies the extent to which data use is valued. Moreover, organizational routines remove the focus from unique occurrences and direct the focus to standard ways of doing things at the school. This provides insight into how data are currently used in practice and how efforts to increase data use might turn out over time. Finally, studying organizational routines results in a more nuanced image of how practices change or persist over time. No one decision, formal structure or person is responsible for this; they all mutually influence each other (Spillane, 2012). This mutual influence comes across in the two aspects of which an organizational routine consists: the ostensive aspect and the performative aspect (Feldman & Pentland, 2003).

2.3. The ostensive aspect

The ostensive aspect of an organizational routine is the schematic, abstract idea of the routine (Feldman & Pentland, 2003), relying on standard operating procedures and taken-for-granted norms. According to Feldman and Pentland, it also includes one's subjective understanding of the routine. These routines never include specific performances, because it is impossible to specify a routine in the amount of detail required to actually carry it out (Blau, 1955; Feldman & Pentland, 2003). An example of the ostensive aspect would be the school principal's vision for working with the data team intervention, as written in the school's policy document. Studying ostensive aspects of organizational routines

illuminates 'how the formal organizational structure is arranged so as to enable and constrain the use and production of various sorts of data' (Spillane, 2012, pp. 124).

2.4. The performative aspect

The performative aspect of an organizational routine captures the everyday practice of specific actions, taken by specific people, at specific times (Feldman & Pentland, 2003). This aspect reflects the knowledge, experience and reflections required to actually take action, whereas the ostensive aspect only offers broad guidance (Spillane, 2012). The performative aspect is inherently improvisational (Feldman & Pentland, 2003), as actions need to be adjusted to changing contexts. In contrast to the example of the ostensive aspect given above, an example of the performative aspect would be the way in which data team members actually work with the data team intervention. Studying these performative aspects of organizational routines helps us to understand how practice unfolds over time (Spillane, 2012). Taken together, the development of the ostensive and performative aspects of organizational routines illustrates the extent to which certain behavior is sustained. But, what types of behavior can be displayed and addressed by schools' organizational routines that aim to sustain data use?

2.5. Data use displayed through organizational routines

Data-based decision making refers to the collection and organization of data that are subsequently used to help improve the quality of the education provided by the individual teachers, schools, or districts (Lai & Schildkamp, 2013). These data can represent aspects of students, teachers, parents, and/or schools, and can be both quantitative and qualitative. Examples are students' achievement scores and observations of classroom teaching. Even though data use is associated with various benefits, previous studies have shown that most teachers do not use data to its best effect or do not use data at all (e.g., Means et al., 2011; Oláh et al., 2010). Therefore, professional development regarding the use of data is urgently needed (Desimone, 2009). The data team intervention is an example of such a professional development program.

Data team intervention. Data teams consist of six to eight people from the same school: 1–2 school leaders, 4–6 teachers, and, if possible, a quality manager, who often has access to data (Schildkamp et al., 2015). Under the supervision of a data coach from the university, team members collaboratively learn how to use data to solve an educational problem at their school. They often choose to work on a problem that is the Inspectorate's point of concern (e.g., high grade retention rates), because Dutch legislation holds schools increasingly accountable for ensuring their educational quality. While working on that problem, they use a structured cyclic procedure as illustrated in Fig. 1, which includes an extensive set of guidelines and activities. Although presented as a linear process, the procedure is more iterative in practice. For example, data team members go back and forth between steps 3, 4 and 5 when the collection of additional data is required.

Sustaining data use through the data team intervention. To determine whether schools who participated in the data team intervention sustain their data use, four sub-behaviors were studied. First, continued engagement in the data team intervention means that both a context and an approach for data use are made available. Such an intervention can predefine and structure the use of data by implementing protocols or procedures which can guide discussions on data use (Gallimore, Ermeling, Saunders, & Goldenberg, 2009). As a result, access to new knowledge is provided.

The second type of behavior that can be displayed to sustain data use is the implementation and evaluation of actions for improvement as designed by the data team members. Important issues can be uncovered and addressed through the use of data while working with the data team intervention (Schildkamp & Poortman, 2015; Schildkamp et al., 2015). For example, when data analysis indicates that students lack proper study skills, teachers can decide to provide a course in which such skills are addressed. These actions for improvement can then be evaluated, for example, to determine their effectiveness. This implies that actions for improvement are both a result of data use and an impetus to collect additional data.

Furthermore, engagement in the data team intervention can stimulate data use for school improvement in general, which has two distinguishable subtypes (Breiter & Light, 2006; Coburn & Talbert, 2006; Schildkamp & Kuiper, 2010; Wohlstetter, Datnow, & Park, 2008). The first is data use for *school development*. Here, data are used as a tool to determine effective teaching methods and professional development needs, and to provide direction for policy development. An example of this is using student results to identify gaps in the curriculum. The second is data use to improve teachers' *instruction*, which is an important factor influencing student achievement (Hattie, 2009). For example, teachers can determine what topics their students are mastering, whether the content is being addressed at the appropriate pace, and what progress is made by individual students. These two types of data use are interrelated. For example, data that are being used to identify gaps in the curriculum can also be used by individual teachers to modify their instruction.

For each of these four behaviors, the development of the ostensive and performative aspects of organizational routines needs to be determined. These behaviors indicate sustainability (or lack of sustainability) when the rigor or frequency of the behavior changes over time. Each behavior should not necessarily be displayed at each school. However, where more types of behaviors are displayed, it seems more likely that sustainable changes were made.

To summarize, the present study aimed to investigate whether schools who implemented the data team intervention sustained their data use. Therefore, the following research question is posed: To what extent is data use sustained in schools using the data team intervention, where sustainability of data use is defined in relation to the extent and manner of development over time of the ostensive and performative aspects of schools' organizational routines regarding: 1) engaging in the data team intervention, 2) acting upon their data team's improvement plan, and using data for 3) school development and 4) instruction?

3. Method

A longitudinal mixed-methods exploratory case study design was used because it can provide a holistic overview of a phenomenon (Gummesson, 1991). Although this means that the statistical generalizability of these findings might be limited (Johnson, 1994), it does provide a deeper insight through analytical generalizability (Yin, 2003).

3.1. Context

Educational setting. This study took place in the Netherlands. In 2012–2013, only 8% of the secondary schools studied by the Dutch Inspectorate used data (Dutch Inspectorate of Education, 2014). The Dutch Ministry of Education believes in the importance of data use for improving education, and stated that the percentage of secondary schools using data needs to increase to at least 90% by 2018 (Verbeek & Odenthal, 2014). Because of this,

The data team members:

1. Define an educational problem on which they want to focus their efforts.
2. Formulate a hypothesis as to the cause of the educational problem.
3. Collect data to investigate the hypothesis.
4. Determine the validity and reliability of the data.
5. Analyze the data (e.g., descriptive or correlational analyses).
6. Interpret the outcomes and draw conclusions. Should the hypothesis turn out to be false, a new hypothesis needs to be investigated (back to step 2). Should the hypothesis turn out to be supported, they proceed to step 7.
7. Develop and implement improvement actions.
8. Evaluate the effectiveness of the measures and modify them if necessary (back to step 7). When the problem is solved, a new problem can be defined (back to step 1).

Fig. 1. The data team intervention (Schildkamp & Ehren, 2013, pp. 56).

schools increasingly prioritize data use.

Dutch schools' organizational routines can vary substantially, as schools have considerable freedom in determining what subject matter they teach and what textbooks, assessments and instructional strategies they use (Kuiper, Van den Akker, Hooghoff, & Letschert, 2006). For example, schools are not obligated to document their policy for each successive school year. Furthermore, they can vary in the way and extent to which they schedule staff meetings. This also implies that schools and data team members have considerable freedom in designing actions for improvement. These measures can vary from changing the school's grade retention policy to changing instructional strategies or textbooks.

Participants. For this project funded by the Dutch Ministry of Education, ten data teams were established in the same number of secondary schools. These schools and their data team members voluntarily participated. As time is an important influencing factor

Table 1
Characteristics of the six participating schools.

School[c]	Educational level[d]	School size: *N* students	Relative size[a]	Denomination[e]	*N* Data team members at time 1	*N* colleagues at time 1
Fairview 1&2	Low	543/464[b]	−	Religious	4/3[b]	43/37
Highland	Interm.-High	1079	=	Public	9	71
Lincoln	Low-High	1261	+	Religious	6	96
Oak Grove	Low-High	1185	+	Religious	6	81
Pine Grove	Interm.-High	1007	−	Religious	6	71
Roosevelt	Low-High	964	−	Public	6	69

[a] Indicates relative school size for its educational level. + 'above average', = 'average', - 'below average'.
[b] Indicates data for locations Fairview1 and Fairview2, respectively.
[c] School and participant names are pseudonyms.
[d] In the Netherlands, there are three main educational tracks. The lowest track (4 years) prepares students for vocational education. The intermediate track (5 years) prepares students for college/higher education. The highest track (6 years) prepares students for university education. These tracks are quite segregated, which means that most students who enter a track stay within that track until graduation.
[e] A school is considered public when it is not affiliated with a religious denomination.

in teachers' professional development, schools that participated needed to revise data team members' work schedule so that participation time was facilitated. Each data team was supervised by the same data coach, who was not an author of this paper. On average, team members had a meeting twice a month for two years. The present study focused on the development of organizational routines during these two years and in the year after, when support from the university was almost entirely withdrawn.[1] Thus, all data teams were studied from their initial set-up up until three years later. Moreover, not only the data team members participated, but also their colleagues (teachers).

At four schools, the response rates for the questionnaire of data team members' colleagues (see Instruments section for more information about the questionnaire) were too low to include them in the present study. The characteristics of the remaining six schools are described in Table 1. All six data teams chose to focus their efforts on increasing their grade retention rates, which received a poor evaluation from the Dutch Inspectorate.

The six data teams were grouped based on their initial level of data use and the type of actions for improvement they designed, to ensure comparability within each group. Of those groups, Fairview 1 and Lincoln were selected through convenience sampling to deepen our insights. In both schools, the principal, two data team members, and two colleagues who did not participate in the data team were interviewed. They were selected to ensure that information was obtained from at least two educators who were familiar with the policy on data use (school leader and principal), at least three (the teachers) who were familiar with everyday teaching practice, and at least two who were familiar with the data team (two team members), and that the two teachers who were less involved in the data team intervention could describe how the intervention influenced their educational practice. See Table 2 for additional information on the interviewees.

3.2. Instruments

Both quantitative and qualitative data were used for the present study. Each instrument was used to collect information about how data use for school development and for instruction were represented in both the ostensive and the performative aspects of organizational routines. The policy documents and interviews were also used to collect information about how engagement in the data team intervention and acting upon the data team's improvement plan were represented in the ostensive and performative aspects of organizational routines.

Questionnaire. Teachers from the six schools participated in a survey three times: when they started working with the data team intervention, after two years of working with it, and one year after support had been almost entirely withdrawn.[1] This questionnaire contained three scales about data use from Schildkamp, Poortman, Luyten, and Ebbeler (2017): one representing schools' overall vision and norms for data use (an ostensive aspect, including both school development and instruction), and scales representing schools' everyday data use practices for school development and instruction (performative aspects).

For each statement from the vision and norms scale and the data use for school development scale, participants indicated the extent to which it applied to their school or their own practice, ranging from *strongly disagree* (1) to *strongly agree* (4). For each statement from the data use for instruction scale, participants' answers concerned frequency, ranging from *never* (1) to *a couple of times per week* (6). Cronbach's alpha for each of the three scales was high, see Table 3.

Policy documents. Information on the new behaviors was also gathered by collecting their policy documents in which they reflected on a calendar year (e.g., 2014), for each year in which the project ran. When the information referred to, for example, a policy, vision or norm, it reflected the ostensive aspect. When the information referred to, for example, a reflection on how data use activities were performed in the everyday practice, it reflected the performative aspect. These documents were collected from all schools except Fairview 1 and 2, as they did not preserve those documents.

Interviews. Semi-structured interviews were conducted at the end of the third year with educators from Fairview 1 and Lincoln, and had an average duration of 45 minutes. Questions were asked about the ostensive and performative aspect of the organizational routines for the four new behaviors (engaging in the data team intervention, acting upon their data team's improvement plan, and using data for school development/instruction). Examples include: In your opinion, what is the school's vision for data use for instruction (ostensive aspect – data use for instruction)? To what extent and in what ways do you use data to improve the education you provide (performative aspect – data use for instruction)? All interviews were audio-recorded and transcribed verbatim.

3.3. Data analysis

Questionnaires, policy documents and interviews were used at multiple time points to determine the development of the schools' organizational routines. This use of multiple sources and methods helped to triangulate the findings and obtain a more complete

[1] Schools could decide to consult the data coach in the third year for up to a maximum of 5 h total. Only Lincoln did so.

Table 2
Characteristics of participants from the case study schools.

	Data team member	Years of experience	Function	Subject
Fairview 1				
Mr. Anderson	Yes	22	Teacher, school leader	Dutch
Mr. Johnson	Yes[a]	14	Principal	–
Ms. Smith	Yes	11	Teacher	English
Mr. White	No	43	Teacher	French
Ms. Thompson	No	7	Teacher	Science & Physics
Lincoln				
Mr. Clark	No	35	Principal	–
Mr. Lewis	No	14	Teacher	Biology
Ms. Harris	Yes	32	School leader	–
Ms. Wright	No	7	Teacher	Management & Organization
Ms. Young	Yes	22	Teacher	Mathematics

[a] Mr. Johnson participated in the data team only during the first and second years.

Table 3
Overview of the Questionnaire's scales.

	Scale	N items	α	Sample item
Ostensive aspect	Overall vision & norms data use (both school development & instruction)	5	0.82	Teachers in my school share a common understanding about effective ways to evaluate student learning
Performative aspect	Data use for school development	10	0.86	Students' results are used to evaluate teachers.
	Data use for instruction	11	0.89	How often do you use data to set learning goals/targets for individual students
Mean response rate at time 1				43.0%
Mean response rate at time 2				33.5%
Mean response rate at time 3				45.2%

picture of the development of the schools' organizational routines. This ensured the construct validity of our findings (Yin, 2003).

Questionnaire. As organizational routines are an aspect of the school's organizational structure (Spillane, 2012), individual answers needed to be aggregated at the school level. Due to privacy issues, questionnaire data could not be linked to specific individuals over time. Thus, data could not be used for t-tests or multi-level analyses. Therefore, all scores were displayed in graphs, which visualized their development over time. After that, data were analyzed in a descriptive manner. As the between-school variation was relatively small for responses on the questionnaire, as well as the policy documents and interviews (see below), results were discussed together.

Policy documents & interviews. A coding scheme was developed, which was inspired by the work of Feldman and Pentland (2003). The coding scheme, presented in Table 4, encompassed the two aspects of organizational routines (ostensive and performative) and the new behaviors (engaging in the data team intervention, acting upon their data team's improvement plan, and using data for school development/instruction). Two codes were applied in Atlas.ti per segment of the policy document or interview: one for the aspect and one for the type of behavior. After that, coded segments were grouped and summarized. For each school, a within-case analysis of the summaries was conducted to describe the development of the organizational routines over time. After that, a cross-case analysis was conducted. It appeared that these descriptions were relatively similar across schools, and they were therefore discussed simultaneously in the results section. Excerpts from the policy documents and quotes from the interviews were translated from Dutch into English to illustrate the results.

To determine inter-rater reliability, 20% of the policy documents (3 out of 15) and the interviews (2 out of 10) were coded by another researcher. This resulted in an inter-rater reliability of 0.74 for the aspect and 0.79 for the behavior, which is considered substantial (Eggen & Sanders, 1993).

4. Results

The results for all six schools are described together, for each of the four behaviors: engaging in the data team intervention, acting upon the data team's improvement plan, and using data for school development and instruction. Policy documents and interviews were used to study all four behaviors, and questionnaire data were also used for data use for school development and instruction. For each behavior, the ostensive aspect of the organizational routines will be addressed first, after which the performative aspect will be described. The results are summarized in Table 5.

4.1. Engaging in the data team intervention

First, it was studied whether schools continued their engagement in the data team intervention, as this can provide both a context and an approach for data use.

Ostensive. The data team intervention was not mentioned in Lincoln and Pine Grove's policy documents and there appeared to be no clear policy or vision for working with data teams. Lincoln's principal Clark was not convinced by the value of team members' work, as he stated during the interview at the end of the third year that: '*When we do something that is some sort of luxury, why do we need it? I would say pretty soon, just cut [the data team intervention] out again.*' The other interviewees from Lincoln did not know what the schools' vision for the data team was. Fairview 1's principal Johnson had no clear vision for the data team intervention either, even though he had been a part of the data team during the first and second years. He explained that *'We let go of the data team intervention* ', but he believed that it is important to use it when a new educational problem arises. The other interviewees from Fairview did not know what principal Johnson's vision for working

Table 4
Coding scheme for policy documents and interviews.

		Ostensive	Performative
		Includes rules, standard operating procedures, norms and plans. Ostensive explains how behaviors should occur, but not how this actually takes place in the educational practice.	Describes how it actually takes place in the educational practice. Concerns specific actions that are (not) taken by specific individuals at specific times.
Engaging in the data team intervention This includes information on the data team intervention and statements on data use interventions in general.		'Our school is going to work with the data team intervention to find an explanation for our high grade retention rates.'	'As a data team, we conducted the following activities: … '
Acting upon the data team's improvement plan Per school, a specific list of actions related to the improvement plan as designed by the data team members (e.g., evaluating students' graduation results) was provided.		'Every year, the school will evaluate students' graduation results per subject. After that, teachers have to make an improvement plan.'	'I made an improvement plan and among other things, I check more regularly whether my students have done their homework.'
Using data for	**School development** Data are used by the school and/or the teachers to determine effective teaching methods, professional development needs, and to provide direction for policy development.	'Data can be a reason to speak with teachers, for example, about the grade retention rates.'	'We conducted analyses and based on those analyses, we decided to pay attention to the pedagogical climate.'
	Instruction Data are used by individual teachers, sometimes together with his/her school leader, to adapt their teaching.	'Teachers should use learning analytics to adapt their education to individual learning needs.'	'After each test, I check for which parts my students score lowest on, and I modify my instruction accordingly.'

Table 5
Summary of results.

Behavior	Ostensive	Performative
Engaging in the data team intervention	The data team intervention was part of schools' policy only to a limited extent. This ostensive aspect remained relatively constant over time.	Five out of six data teams continued to exist in the third year, indicating that the performative aspect remained relatively constant over time. All teams deviated at least to some extent from the original intervention.
Acting upon the data team's improvement plan	None of the schools developed a clear policy on the actions related to the team's improvement plan. This ostensive aspect did not develop over time.	Actions related to the team's improvement plan were not (entirely) being implemented, or implemented in a way that was inconsistent with previous research activities. These actions were not evaluated. This performative aspect did not develop over time.
Using data for school development	All six schools established policy for data use for school development to some extent, but this policy remained relatively constant over time.	Data were used to some extent. This performative aspect did not develop over time.
Using data for instruction	None of the schools developed a clear policy for data use for instruction. This ostensive aspect remained relatively constant over time.	This performative aspect increased slightly over time, from a couple of times per year to monthly.

with data teams was. Moreover, the two teachers were even unsure whether the data teams still existed.

At Highland, statements about the data team intervention and what this intervention roughly entails were found only in their first year's policy document. At Oak Grove, similar statements were found, but only in their second year's policy document. The only school where the policy for the data team intervention was described throughout the years, although in a limited way, was Roosevelt. For example, in the third year, the policy document stated that: '*the data team will continue studying new questions and will receive limited* support'. No further information (e.g., no vision or policy) was provided. Taking all of this together, the data team was part of schools' vision only to a limited extent, as the reasons why the school worked with this intervention were never included. Moreover, this ostensive aspect remained relatively constant over time.

Performative. Even though the ostensive aspect was barely addressed in the policy documents, brief statements on the actual practices of the data team were included in documents from all schools except Lincoln's. Most often, this included a statement about what problem the team had studied. These descriptions were most elaborate in Highland's documents, which also included what hypotheses had been studied. For example, their second year's document stated: *'The hypotheses that have been studied so far*

include the chances of graduation for students who did or did not repeat a grade in the previous years, the influence of our achievement margins within which we discuss whether students need to repeat a grade, and the predictive value of our tests.'

At the end of the third year, it appeared that all data teams except for Pine Grove's had been continued. In some teams, members continued to work on their statement of a problem (grade retention rates), in other teams, new areas of research were chosen. For example, at Lincoln, team members had studied the effect of classroom size on student achievement. Overall, members who had already been participating in the data team continued to do so. Sometimes, one or two new members were included. Spin-off teams were never formed.

All teams deviated at least to some extent from the original data team intervention, as none of the teams evaluated whether they had solved the educational problem. Other examples of deviations include that Lincoln's team members worked in three different subgroups to progress faster, and Fairview 1's members did not use the intervention at all during the third year, because they aligned their tasks with the existing quality management system. They collected data to inform teachers about which students were falling behind.

Interviewees' opinions about the influence of the data team were divided. For example, some believed the data team had helped to create awareness for data use. However, others did not acknowledge this. For example, Fairview 1's principal Johnson (a former data team member) explained: '*The quality managers have had an influence, the data team members have not. We have used the data team intervention to conduct research activities, not to implement data use.*'

Taking all of this together, the performative aspect of the organizational routine was developed to some extent, but remained relatively constant over time. In addition, all teams deviated at least to some extent from the original intervention, as none of the teams evaluated whether they had solved the educational problem, or made other changes in the intervention.

4.2. Acting upon the data team's improvement plan

Second, it was studied how data team members and their colleagues acted upon the team's actions for improvement, which can be both a result of data use and an impetus to collect additional data.

Ostensive. All data teams had used data to design actions for improvement to decrease their grade retention rates. Only Highland's policy documents described the principal's policy for the implementation of the planned actions for improvement. For example, they included the statement that '*within the subject groupings, we collaboratively take care of our tests and assessments. We pay attention to the coherence in the curriculum and the predictive value of grades.*' Additional information about how this should be done was not included. In the other schools, the actions related to the data teams' improvement plan were not described in the policy documents. Moreover, none of the interviewees from Fairview 1and Lincoln could describe the policy for the actions for improvement. Thus, the ostensive aspect of the organizational routines for actions related to the team's improvement plan seemed underdeveloped and remained relatively constant over time.

Performative. All data teams had struggled with implementation, and it appeared that none of the teams had implemented all actions related to the improvement plan at their school. Some had not even implemented any actions at all. This was the case at both Lincoln and Fairview 2. The interviews illuminated that at Lincoln those actions were no longer relevant, as other measures had already been taken. Ms. Harris explained this by stating: '*An organization cannot wait two years before actions are implemented.*'

At Fairview 1, Oak Grove and Highland, the actions related to the improvement plan were implemented partially. For example, at Fairview 1, the data team members made an observation scheme to give teachers feedback on their classroom management role. However, it appeared that this scheme was not yet used by everyone. For example, Ms. Smith noted that she did not yet use the form to observe a colleagues' lesson, and neither did a colleague use it to observe her lesson. In addition, the interviewed teachers from Fairview 1 were unaware that the data team members had designed actions for improvement.

Even though actions for improvement were implemented at Pine Grove and Roosevelt, this was done in ways that were inconsistent with their research activities. For example, at Roosevelt, data team members had concluded that counselors' roles needed to be modified for students at the highest educational level. But their role was also modified for students at the lowest and intermediate educational levels, which was beyond the scope of their study. In no team did members evaluate the way in which the actions related to the improvement plan were implemented.

Taking all of this together, it appeared that the performative aspect remained relatively constant over time. Actions related to the team's improvement plan were not being implemented, were only partially implemented, or were implemented in a way that was inconsistent with team members' research activities. Furthermore, an important part of data use was left out, as the implementation of the actions for improvement was never evaluated.

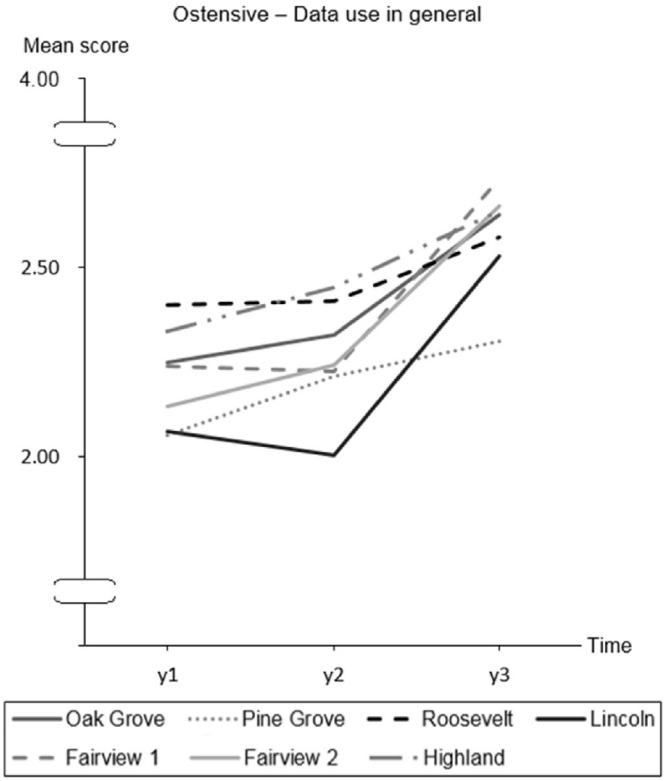

Fig. 2. Development over time of the ostensive aspect of routines related to overall data use (including data use for both school development and data use). Note. The y-axis represents teachers' mean score for this scale, aggregated at the school level. The scale was composed of statements for which scores could range from *strongly disagree* (1) to *strongly agree* (4).

4.3. Using data for school development

Third, it was studied how data was used to determine effective teaching methods and professional development needs, and to provide direction for policy development.

Ostensive. Overall, the questionnaire data show that teachers gave rather similar judgments about their school's overall vision and norms for data use (including data use both for school development and for instruction), see Fig. 2. The scores were a bit below the scale midpoint of 2.5 (where higher scores show stronger agreement), which indicates that teachers were not so sure whether there was a policy. This ostensive aspect changed little over time. The specific development of the ostensive aspect for school development is described next.

Roosevelt was the only school where policy documents did not include a vision for data use for school development. In the other four schools, it was included, albeit in an inconsistent manner. For example, at Pine Grove, data use for school development was included at time points 1 and 2, but not at time point 3. It might be that the principal no longer wanted to use data for school development, but this could not be deduced from the policy documents. The concreteness of principals' vision varied. The vision at Lincoln was least concrete, stating that: '*Besides the fact that schools should have their results in order, we want to improve our results by using a pedagogical approach that puts students at the center.*' No additional information was provided. The vision at Oak Grove was the most concrete, stating that: '*For each course, the school has to map the grades, which will be analyzed. When the differences between our own exams and nation-wide exams are larger than 0.5%, the underlying causes will be determined and actions will be taken to improve this. Per course, teachers have to reach agreement regarding the grading policy*'.

The interviews illustrated that there was a clear vision for data use for school development at Fairview 1. For a few years now, it has been this school's policy that the quality manager presents data about the quality of education on a fixed day of the year. Furthermore, teachers in departments receive data about their own subject. When grades deviate too much from the national average, either higher or lower, teachers have to make a plan for improvement, and evaluate whether this plan worked. They do not receive guidelines about how to write such a plan, but they can ask the quality manager for help. At Lincoln, principal Clark envisioned that data needed to be used by teachers in departments, and that students needed to perform above the national average and be on their way to good or even excellent performance. However, teachers were not obligated to write down their plan of action. Lincoln's teachers Mr. Lewis and Ms. Wright knew that they are expected to use data, but neither of them knew exactly how they are expected to do so. Overall, the ostensive aspect of organizational routines for data use for school development was developed to some extent, but did not change over time.

Performative. None of the documents included information on how teachers had actually used data for school development in practice. The questionnaire responses illustrated that the level of teacher agreement regarding data use for school development was approximately similar within all schools, and was slightly above the scale midpoint of 2.5, see Fig. 3.

Furthermore, interviewees from both Fairview 1and Lincoln stated that not all departments actually used data to design plans for improvement. Even when those plans were designed, they were neither executed nor evaluated. For example, Lincoln's teacher Ms. Wright stated that: '*We see the data, but we do not act upon it.*' At Fairview 1, most interviewees indicated that this type of data use was not being monitored by the school leaders, whereas they deemed this to be necessary. For example, teacher Ms. Smith stated: '*It is a pity, those plans are dropped on the principal's desk, and he needs to take actions and check whether those plans are being executed. And that is not being done*'. Thus, this type of data use did not go beyond the mere signaling of problems. Overall, the performative aspect of organizational routines for data use for school development was developed to a limited extent and changed little over time.

4.4. Using data for instruction

Finally, it was studied how data was used by individual teachers to modify their instruction.

Ostensive. Data use for instruction was absent in all policy documents, except in Highlands' document at time 3. In this document, it said '*Data use: teachers already recognize differences in ability level between students. However, it is not yet self-evident that they use this information to adapt their instruction and assignments.*'

At both Fairview 1 and Lincoln, there did not appear to be a policy for how teachers should use data to improve their own functioning. However, the teachers at Fairview 1 who were interviewed felt that they were being held accountable for their students' achievement. For example, Ms. Thompson stated: '*I do not agree with it, but within this school, there seems to be a policy that your classroom average has to be around 6.7 […] I think this is not fair, but sometimes, it is being expected by our school leaders.*' Overall, the ostensive aspect of organizational routines for data use for instruction was underdeveloped and did not change over time.

Performative. None of the documents included information on how teachers had actually used data for instruction in their practice. Results from the questionnaire illustrated that the average frequency of data use for instruction was approximately similar within all schools, and was below the scale midpoint of 3.5 at time points 1 and 2, and slightly above it at time point 3, see Fig. 3. This indicates that during the first two years, teachers used data to inform their instruction on average a few times per year, and during the third year, this had increased to a maximum average of once a month (scores of 3 and 4 were labeled as "a couple of times per year" and "monthly", respectively).

The three interviewed teachers from Lincoln all used data to improve their instruction in practice. For example, Ms. Wright divided her tests into a language and a mathematics component, so she could tell students exactly which component they had not yet mastered. Unfortunately, the accountability pressure felt by Fairview 1's teachers caused them to pull down or inflate students' grades (depending on whether students' grades were considered to be too high or too low, respectively). For example, Mr. White deliberately included difficult open-ended questions in his vocabulary tests to pull down students' grades, and Ms. Smith modified her grading policy so that which students' grades became higher. The principal and the school leader were not aware of this pressure being felt by teachers. Overall, the performative aspect of organizational routines for data use for instruction was developed to some extent and slightly increased over time.

5. Conclusions and discussion

Professional development and school improvement do not automatically take place as a result of working in professional learning communities. The present study aimed to investigate the sustainability of schools' data use by studying the development over time of the ostensive and performative aspects of schools' organizational routines regarding: 1) engaging in the data team intervention (which is an example of a professional learning community), 2) acting upon the data team's improvement plan, and using data for 3) school development and 4) instruction. The

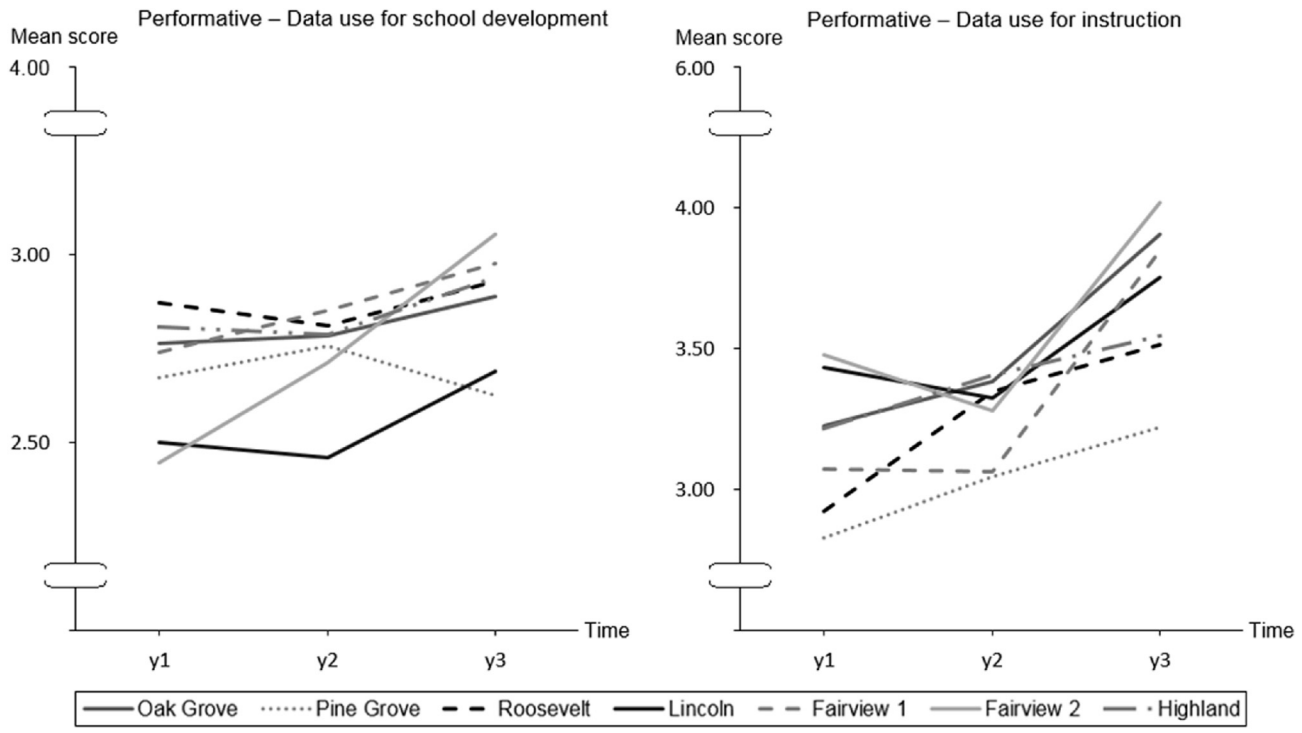

Fig. 3. Development over time of the performative aspect of routines related to data use for school development and instruction. Note. The y-axis represents teachers' mean score for the scale, aggregated at the school level. The "school development" scale was composed of statements for which scores could range from *strongly disagree* (1) to *strongly agree* (4). For "instruction" scale, scores could rage from *never* (1) to *a couple of times per week* (6).

conclusions will be described separately for the ostensive and the performative aspect and organizational routines in general.

5.1. The ostensive aspect

Within the schools, there was only limited policy and vision related to engaging in the data team intervention and using data for school development. Furthermore, it appeared that a policy or vision for acting upon the data team's improvement plan and using data for instruction was lacking. In all schools, the ostensive aspect remained more or less similar over time. Taken together, the lack of vision and policy is likely to have impeded schools' sustained data use, as teachers were not provided with clear guidelines on how to use data. In that way, the enactment of the performative aspect must rely heavily on the teachers' own capacity to use data, which could result in unstable (e.g., teachers only use data when they 'feel like it') and diffuse (e.g., teachers do not use data in a uniform way) practices. Such data use practices are unlikely to result in substantial school improvement. Therefore, it is essential that school leaders provide active support for data use (e.g., Farley-Ripple & Buttram, 2014; Levin & Datnow, 2012; Marsh & Farrell, 2015). For example, establishing a clear school-wide vision is critical for implementing data use, because it can increase teachers' buy-in or belief in the importance of data use (Schildkamp & Kuiper, 2010; Verhaege, Vanhoof, Valcke, & Van Petegem, 2010). This agrees with previous research (e.g., Dede & Rockman, 2007; Desimone, 2002) that factors such as strong school leadership and establishing shared goals are critical to facilitate the implementation of school-wide change. This also signals the need to determine the specific school-context regarding the (perceived) policy for data use. Overall, a stronger focus on the development of the ostensive aspect seems required.

5.2. The performative aspect

Within the schools, engagement in the data team intervention, acting upon the data team's improvement plan, and using data for school development occurred to some extent in everyday practice. However, this performative aspect remained relatively constant over time. It appeared that five out of the six data teams continued their activities. Overall, the members who had already been participating in the data team continued their participation, but spin-off teams were never set up. Strikingly, all data teams deviated at least to some extent from the original intervention, as none of them evaluated whether the educational problem was actually solved. As will be argued later on, the lack of evaluation is likely to have a negative impact on the quality of data use. Other deviations, such as working in small sub-groups, might be less consequential.

Concerning the actions related to the improvement plan designed by the data team members, it appeared that these actions were not being implemented, were only partially implemented, or were implemented in a way that was inconsistent with data team members' research activities. Regarding the use of data for school development, it appeared that data were only used to signal problems. Educators did not act upon these data.

Interestingly, educators never evaluated whether the educational problem was solved, whether the actions for improvement as designed by their data team were implemented accurately, or whether their data use for school development improved the quality of education. As a result, an essential aspect of data use was omitted (Bryk, Gomez, Grunow, & LeMahieu, 2015; Mandinach, Honey, Light, & Brunner, 2008; Marsh, 2012), for example, because one should anticipate and measure any unintended effects of their actions (Bryk et al., 2015). As the cycles of data use are completed only after evaluation takes place and the educational problem is solved, the ways in which data were used in data teams

and schools can at best be described as being incomplete. One of the reasons why data team members struggled to engage in implementation and evaluation activities might have been the withdrawal of the coach's support. During the first and the second year, team members had, on average, a meeting twice a month. The coach supported their data use during those meetings. In the third year, the support was brought down to an optional 5 hours of support, of which only Lincoln made use. It might be that the withdrawal of this support was too sudden and too extensive, and that data team members did not yet had enough knowledge and skills to pursue the implementation and evaluation aspects of the data use cycle.

Regarding data use for instruction, it appeared that teachers slightly increased this type of data use over time. This indicated that over time, the performative aspect had developed at least to some extent. However, sometimes this resulted in misuse. For example, at Fairview 1, teachers felt pressured to meet certain benchmarks. As a result, they pulled down or inflated students' grades, which decreases the validity and reliability of students' test results. This echoes previous research illustrating that (the perception of) organizational routines can contribute to mindless action and inappropriate responses to problems (e.g., Levitt & March 1988). Such accountability pressures are undesirable and should be targeted more clearly. Moreover, the pressure felt at Fairview 1 signals that the specific school-context could play an important role as well.

Overall, the present study illustrated that enactment on the performative aspect of data use was likely to rely heavily on teachers' own capacity to use data, as the ostensive aspect was relatively underdeveloped. Moreover, it might be that teachers' attitude towards data use influenced their (lack of) data use. For example, previous research illustrated that teachers' perceptions of conditions for workplace learning and attitudes toward reform in terms of learning and teaching are mutually related, and influence the ways in which teachers enact in the reform (Imants, Wubbels, & Vermunt, 2013). In addition, previous research illustrated that organizational responsibilities and reform history shape how individuals come to understand data use practices (Coburn & Talbert, 2006).

5.3. The development of organizational routines

First, the present study illustrated that organizational routines were scarce and sometimes even lacking. Furthermore, these routines developed little over time. Thus, even though it was hypothesized that working with the data team intervention could result in lasting school-wide changes, the findings illustrated this was not (yet) the case. This might have been caused by data team members' struggle to broker their knowledge about data use and the actions resulting from their improvement plan to their colleagues (Hubers et al., 2017). If so, an important opportunity to facilitate school staff's participation in discussions of school-wide issues and to increase communication about data use was missed. Given the importance of leadership, vision and policy for implementing data use (Levin & Datnow, 2012), one way to stimulate the development of organizational routines might be to focus on developing the ostensive aspect (Sherer & Spillane, 2011; Spillane, Parise, & Sherer, 2011). It might also be that the sustainability of schools' data use was investigated too early, as it can take five to ten years before a school is completely reformed (Desimone, 2002).

Second, the present study illustrated the importance of distinguishing both the ostensive and performative aspect of organizational routines. This combination offered a complete and thorough insight into the 'status quo' at each school. This agrees with Spillane (2012), who argued that the combination of both aspects is essential in studying the issue of sustainability. Moreover, it appeared to be critical to study organizational routines over time to determine whether a certain level of data use reflected sustained school-wide change. This also agrees with Spillane (2012).

5.4. Practical implications

The present study illustrated that more attention should be paid to the development of the ostensive aspect of organizational routines for data use, as both data use for school development and data use for instruction changed little over time. This agrees with Gerzon (2015), who indicated that school leaders should nurture a culture for data use, for example by providing resources and communicating expectations for data use more clearly. This might be done by including the principal in the data team. Moreover, team members could be involved in the development of guidelines, agendas or discussion protocols with guiding questions for effective data use (Blanc et al., 2010). Furthermore, it might also be necessary that principals monitor the extent to which teachers actually use data, and continuously determine how best to support them. Moreover, to aid the development of the performative aspect of organizational routines for data use, attention should be paid to teachers' capacities to use data. Even though previous research illustrated that data team members' gained knowledge about data use after participating in the data team intervention, it might be that such results are enhanced when additional attention is paid to individual factors (e.g., attitude towards data, teachers' beliefs, individual's sensemaking and agency; Coburn & Talbert, 2006; Higgins & Spitulnik, 2008). In addition, attention should be paid to the way in which data team members broker their knowledge to their colleagues in order to instruct colleagues' data use. For example, previous research illustrated that data team members struggled greatly to do this, and used both few and superficial manners to do so (Hubers et al., 2017; Hubers, Poortman, Schildkamp, & Pieters, submitted). Moreover, even when they brokered their knowledge, it was almost always knowledge about the educational problem, and not knowledge about data use itself.

These practical implications might have a wider scope than just the data use context, as previous research has illustrated that one of the biggest challenges of professional learning communities is to sustain teachers' professional development (e.g., Harris & Jones, 2010; Van Veen, Zwart, Meirink, & Verloop, 2010). It might be that educators in other such bottom-up and small-scale professional development programs face similar challenges with the development of organizational routines. In that case, the aforementioned practical implications can be a starting point in providing wider support for sustaining professional development initiatives.

5.5. Limitations

The present study had a few limitations. First, questionnaire data could not be linked over time, as identifiers were not included due to privacy issues. This prohibited us from conducting certain statistical analyses which might have provided more fine-grained results. But, as the graphs hardly display change over time, such analysis would have been likely to point out the non-significance of the changes in schools' routines over time. Future research should try to include multilevel designs. However, as a sample size of at least 30 groups with 30 individuals per group is required (Hox, 2010), with even a higher number of groups required for more sophisticated analyses, this might be a challenge for educational researchers.

Second, according to Spillane (2012), conducting observations is important for understanding how practice unfolds over time.

Moreover, it can give a more objective description of change over time than self-reports (Van Driel, Meirink, Van Veen, & Zwart, 2012). However, as educators from the present study indicated that they did not increase their data use over time and reported struggles like accountability pressure, we have no clear reason to believe they were too positive. Moreover, we triangulated data from questionnaires, policy documents and interviews to describe the organizational routines and corroborate our findings.

Third, we did not include instruments to gain insight into the quality of the work with the data team intervention and the data use. It could be that the occurrence of data use did not vary between schools, but the quality did. Even though quality issues came up (e.g., all teams deviated from the original data team intervention), it might be that additional issues were present.

5.6. Future research

Future research should continue to study how professional development programs result in sustained school-wide changes. To do so, more insight needs to be gained into the development of organizational routines. One important area of research is to determine how ostensive and performative aspects influence each other, and how the development of both can be stimulated. What principals can do to not only establish a clear, widespread policy, but also to facilitate a context in which teachers actually enact that policy needs to be determined.

Furthermore, additional insight needs to be gained in reasons why professional development programs do (not) result in sustained school-wide changes. The reason for this is that successful school improvement is context-specific and is closely related to schools' capacity to change their own context (Sleegers & Leithwood, 2010). Frameworks such as 'realistic evaluation' and 'theory of change' (Pawson & Tilley, 1997; Weiss, 1998) might be helpful in that regard, as those focus on what works for whom and in what context.

Another important issue for future research is to include measures of fidelity, which reflects the degree to which an intervention is carried out as intended (Keller-Margulis, 2012). Through including such information, clearer conclusions can be drawn about the sustainability of an intervention. For example, if routines for data use are developed, but low fidelity to the program is displayed (e.g., evaluation did not take place), it is unlikely that the development of routines is caused by the program.

Even though several questions remain to be addressed, studying the development of organizational routines can be considered a fruitful effort to understand whether and how professional development programs result in sustainable school-wide change (e.g., Coburn & Turner, 2011; Little, 2012; Spillane, 2012). This will allow for informed decisions as to whether or not a program is worth the investment of efforts and resources (Coburn & Turner, 2011). The present study illustrated that sustaining data use is indeed a quest: schools were not (yet) able to develop organizational routines around data use. Therefore, interventions should target such development more clearly to support schools in sustaining their data use.

Acknowledgement

The authors are grateful to the Dutch Ministry of Education, Culture & Science for funding the data team project. Although government-funded, the present study was independently designed and conducted.

References

Blanc, S., Christman, J. B., Liu, R., Mitchell, C., Travers, E., & Bulkley, K. E. (2010). Learning to learn from data: Benchmarks and instructional communities. *Peabody Journal of Education, 85*(2), 205–225. http://dx.doi.org/10.1080/01619561003685379.

Blau, P. (1955). *The dynamics of bureaucracy*. Chicago: University of Chicago Press.

Breiter, A., & Light, D. (2006). Data for school improvement: Factors for designing effective information systems to support decision-making in schools. *Educational Technology & Society, 9*(3), 206–217.

Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015). *Learning to improve: How America's schools can't get better at getting better*. Cambridge, MA: Harvard Education Press.

Campbell, C., & Fullan, M. (2006). *Unlocking the potential for district wide reform*. Paper available at: http://www.michaelfullan.ca/Articles_06/Articles_06a.htm Accessed on 12 May 2015.

Carlson, D., Borman, G., & Robinson, M. (2011). A multistate district-level cluster randomized trial of the impact of data-driven reform on reading and mathematics achievement. *Educational Evaluation and Policy Analysis, 33*(3), 378–398.

Chapman, C., & Muijs, D. (2014). Does school-to-school collaboration promote school improvement? A study of the impact of school federations on student outcomes. *School Effectiveness and School Improvement, 25*(3), 351–393.

Coburn, C. E., & Russell, J. L. (2008). District policy and teachers' social networks. *Educational Evaluation and Policy Analysis, 30*(3), 203–235.

Coburn, C. E., & Talbert, J. E. (2006). Conceptions of evidence use in school districts: Mapping the terrain. *American Journal of Education, 112*, 469–495.

Coburn, C. E., & Turner, E. O. (2011). Research on data use: A framework and analysis. *Measurement, 9*, 173–206.

Crossan, M. M., & Apaydin, M. (2010). A multi-dimensional framework of organizational innovation: A systematic review of the literature. *Journal of Management Studies, 47*(6), 1154–1191. http://dx.doi.org/10.1111/j.1467-6486.2009.00880.x.

Datnow, A., & Hubbard, L. (2015). Teachers' use of assessment data to inform instruction: Lessons from the past and prospects for the future. *Teachers College Record, 117*(4), 1–26.

Datnow, A., Park, V., & Kennedy-Lewis, B. (2013). Affordances and constraints in the context of teacher collaboration for the purpose of data use. *Journal of Educational Administration, 51*(3), 341–362.

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.

Dede, C., & Rockman, S. (2007). Lessons learned from studying how innovations can achieve scale. *Threshold*, 4–10.

Desimone, L. M. (2002). How can comprehensive school reform models be successfully implemented? *Review of Educational Research, 72*(3), 433–479. http://dx.doi.org/10.3102/00346543072003433.

Desimone. (2009). Improving impact studies of teacher's professional development: Toward better conceptualizations and measures. *Educational Researcher, 38*(3), 181–199.

Van Driel, J. H., Meirink, J. A., Van Veen, K., & Zwart, R. C. (2012). Current trends and missing links in studies on teacher professional development in science education: A review of design features and quality of research. *Studies in Science Education, 48*(2), 129–160. http://dx.doi.org/10.1080/03057267.2012.738020.

Dutch Inspectorate of Education. (2014). *De staat van het onderwijs* [The State of Affairs in Education]. Utrecht: Inspectie van het Onderwijs.

Ebbeler, J., Poortman, C. L., Schildkamp, K., & Pieters, J. M. (2016). Effects of a data use intervention on educators' use of knowledge and skills. *Studies in Educational Evaluation, 48*, 19–31. http://dx.doi.org/10.1016/j.stueduc.2015.11.002.

Eggen, T. J. H. M., & Sanders, P. F. (1993). *Psychometrie in de Praktijk* [Psychometrics in Practice]. Arnhem: CITO.

Farley-Ripple, E. N., & Buttram, J. L. (2014). Developing collaborative data use through professional learning communities: Early lessons from Delaware. *Studies in Educational Evaluation, 42*, 41–53. http://dx.doi.org/10.1016/j.stueduc.2013.09.006.

Feldman, M. S., & Pentland, G. T. (2003). Reconceptualizing organizational routines as a source of flexibility and change. *Administrative Science Quarterly, 48*, 94–118.

Fullan, M. (2007). *The new meaning of educational change*. New York, NY: Teachers College Press.

Gallimore, R., Ermeling, B. A., Saunders, B., & Goldenberg, C. (2009). Moving the learning of teaching closer to practice: Teacher education implications of school-based inquiry teams. *Elementary School Journal, 109*(5), 537–553. http://dx.doi.org/10.1086/597001.

Gerzon, N. (2015). Structuring professional learning to develop a culture of data use: Aligning knowledge from the field and research findings. *Teachers College Record, 117*(4), 1–28.

Gummesson, E. (1991). *Qualitative methods in management research*. California: Sage Publication.

Harris, A., & Jones, M. (2010). Professional learning communities and system improvement. *Improving Schools, 13*(2), 172–181.

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

Higgins, T. E., & Spitulnik, M. W. (2008). Supporting teachers' use of technology in science instruction through professional development: A literature review.

Journal of Science Education and Technology, 17, 511–521. http://dx.doi.org/10.1007/s10956-008-9118-2.
Honig, M. I. (2008). District central offices as learning organizations: How sociocultural and organizational learning theories elaborate district central office administrators' participation in teaching and learning improvement efforts. *American Journal of Education, 114*, 627–664.
Hox, J. J. (2010). *Multilevel analysis: Techniques and applications*. New York: Routledge.
Hubers, M. D., Moolenaar, N. M., Schildkamp, K., Daly, A. J., Handelzalts, A., & Pieters, J. M. (2017). Share and succeed: The development of knowledge sharing and brokerage in data teams' network structures. *Research Papers in Education*. http://dx.doi.org/10.1080/02671522.2017.1286682.
Hubers, M.D., Poortman, C.L., Schildkamp, K., and Pieters, J.M. Spreading the word: Boundary crossers building collective capacity for data use, submitted for peer review.
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., & Honig, M. I. (2010). Tools to deepen practitioners' engagement with research: The case of the Institute for Learning. In C. E. Coburn, & M. K. Stein (Eds.), *Research and practice in education: Building alliances, bridging the divide*. New York, NY: Rowman & Littlefield.
Imants, J., Wubbels, T., & Vermunt, J. D. (2013). Teacher's enactments of workplace conditions and their beliefs and attitudes toward reform. *Vocations and Learning, 6*, 323–346. http://dx.doi.org/10.1007/s12186-013-9098-0.
Johnson, D. (1994). *Research methods in educational management*. Essex: Longman Group.
Karr, K. A., Marsh, J. A., Ikemoto, G. S., & Barney, H. (2006). Strategies to promote data use for instructional improvement: Actions, outcomes, and lessons from three urban districts. *American Journal of Education, 112*(4), 496–520.
Keller-Margulis, M. (2012). Fidelity of implementation framework: A critical need for response to intervention models. *Psychology in the Schools, 49*(4), 342–352. http://dx.doi.org/10.1002/pits.21602.
Kuiper, W., Van den Akker, J., Hooghoff, H., & Letschert, J. F. M. (2006). Curriculum policy and school practice in a European comparative perspective. In J. F. M. Letschert (Ed.), *Curriculum development Re-invented. Proceedings of the invitational conference on the occasion of the 30 years SLO 1975-2005* (pp. 56–77). Enschede: SLO.
Kurland, H., Peretz, H., & Hertz-Lazarowitz, R. (2010). Leadership style and organizational learning: The mediate effect of school vision. *Journal of Educational Administration, 48*(1), 7–30. http://dx.doi.org/10.1108/09578231011015395.
Lachat, M. A., & Smith, S. (2005). Practices that support data use in urban high schools. *Journal of Education for Students Placed at Risk, 10*(3), 333–349.
Lai, M. K., & Schildkamp, K. (2013). Data-based decision making: An overview. In K. Schildkamp, M. K. Lai, & L. Earl (Eds.), *Data-based decision making in education: Challenges and opportunities* (pp. 9–21). Dordrecht: Springer.
Levin, J. A., & Datnow, A. (2012). The principal role in data-driven decision making: Using case-study data to develop multi-mediator models of educational reform. *School Effectiveness and School Improvement, 23*(2), 179–201. http://dx.doi.org/10.1080/09243453.2011.599394.
Levitt, B., & March, J. G. (1988). Organizational learning. *Annual Review of Sociology, 14*, 319–340.
Little, J. W. (2012). Understanding data use practice among teachers: The contribution of micro-process studies. *American Journal of Education, 118*(2), 143–166. http://dx.doi.org/10.1086/663271.
Lomos, C., Hofman, R. H., & Bosker, R. J. (2011). Professional communities and student achievement - a meta-analysis. *School Effectiveness and School Improvement, 22*(2), 121–148. http://dx.doi.org/10.1080/09243453.2010.550467.
Mandinach, E., Honey, M., Light, D., & Brunner, C. (2008). A conceptual framework for data-driven decision-making. In E. B. Mandinach, & M. Honey (Eds.), *Data-driven school improvement: Linking data and learning* (pp. 13–31). New York, NY: Teachers College Press.
March, J. G., & Simon, H. A. (1958). *Organizations*. New York: Wiley.
Marsh, J. A. (2012). Interventions promoting educators' use of data: Research insights and gaps. *Teachers College Record, 114*(11), 1–48.
Marsh, J. A., Bertrand, M., & Huguet, A. (2015). Using data to alter instructional practice: The mediating role of coaches and professional learning communities. *Teachers College Record, 117*(4), 1–40.
Marsh, J. A., & Farrell, C. C. (2015). How leaders can support teachers with data-driven decision making: A framework for understanding capacity building. *Educational Management Adaministration & Leadership, 43*(2), 269–289. http://dx.doi.org/10.1177/1741143214537229.
Marsh, J. A., McCombs, J. S., & Martorell, F. (2009). How instructional coaches support data-driven decision making. *Educational Policy, 24*(6), 872–907.
McDougall, D., Saunders, W., & Goldenberg, C. (2007). Inside the black box of school reform: Explaining the how and why of change at getting results schools. *International Journal of Disability, Development and Education, 54*(1), 51–89.
Means, B., Chen, E., DeBarger, A., & Padilla, C. (2011). *Teachers' ability to use data to inform instruction: Challenges and supports*. Washington, DC: U.S. Department of Education, Office of Planning, Evaluation and Policy Development.
Nelson, R. R., & Winter, S. G. (1982). *An evolutionary theory of economic change*. Cambridge, MA: Harvard University Press.
Oláh, L. N., Lawrence, N. R., & Riggan, M. (2010). Learning to learn from benchmark assessment data: How teachers analyze results. *Peabody Journal of Education, 85*(2), 226–245. http://dx.doi.org/10.1080/01619561003688688.
Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. London: SAGE.
Penuel, W. R., Fishman, B. J., Cheng, B. H., & Sabelli, N. (2011). Organizing research and development at the intersection of learning, implementation, and design. *Educational Researcher, 40*(7), 331–337. http://dx.doi.org/10.3102/0013189X11421826.
Poortman, C. L., & Schildkamp, K. (2016). Solving student achievement problems with a data use intervention for teachers. *Teaching and Teacher Education, 60*, 425–433.
Schildkamp, K., & Ehren, M. (2013). The Netherlands: From "intuition"- to "data"-driven decision making in Dutch secondary schools? In K. Schildkamp, M. K. Lai, & L. Earl (Eds.), *Data-based decision making in education: Challenges and opportunities* (pp. 49–68). Dordrecht: Springer.
Schildkamp, K., & Kuiper, W. (2010). Data-informed curriculum reform: Which data, what purposes, and promoting and hindering factors. *Teaching and Teacher Education, 26*(3), 482–496.
Schildkamp, K., & Poortman, C. L. (2015). Factors influencing the functioning of data teams. *Teachers College Record, 117*(4), 1–30.
Schildkamp, K., Poortman, C. L., & Handelzalts, A. (2015). Data teams for school improvement. *School Effectiveness and School Improvement, 27*(2), 228–254. http://dx.doi.org/10.1080/09243453.2015.1056192.
Schildkamp, K., Poortman, C. L., Luyten, H., & Ebbeler, H. (2017). Factors promoting and hindering data-based decision making in schools. *School Effectiveness and School Improvement, 28*(2), 242–258. http://dx.doi.org/10.1080/09243453.2016.1256901.
Sherer, J. Z., & Spillane, J. P. (2011). Constancy and change in work practice in schools: The role of organizational routines. *Teachers College Record, 113*(3), 611–657.
Sleegers, P., & Leithwood, K. (2010). School development for teacher learning and change. In P. Peterson, E. Baker, & B. McGaw (Eds.), *7. International encyclopedia of education* (pp. 557–562). Oxford: Elsevier.
Spillane, J. P. (2012). Data in practice: Conceptualizing the data-based decision-making phenomena. *Teachers College Record, 118*(2), 113–141.
Spillane, J. P., Parise, L. M., & Sherer, J. Z. (2011). Organizational routines as coupling mechanisms: Policy, school administration and the technical core. *American Educational Research Journal, 48*(3), 589–619. http://dx.doi.org/10.3102/0002831210385102.
Van Geel, M., Keuning, T., Visscher, A. J., & Fox, J. P. (2016). Assessing the effects of a school-wide data-based decision making intervention on student achievement growth. *American Educational Research Journal, 53*(2), 360–394.
Van Veen, K., Zwart, R., Meirink, J., & Verloop, N. (2010). Professionele ontwikkeling van leraren. Een reviewstudie naar effectieve kenmerken van professionaliseringsinterventies van leraren. [Teachers' professional development. A review study of effective characteristics of teachers' professional development programs]. ICLON/Expertisecentrum Leren van Docenten.
Verbeek, C., & Odenthal, L. (2014). Opbrengstgericht werken en onderzoeksmatig leiderschap in PO en VO [DBDM and research leadership in primary and secondary education. In M. Krüger (Ed.), *Leidinggeven aan onderzoekende scholen* (pp. 67–78). Bussum: Coutinho: Leading Researching Schools.
Verhaege, G., Vanhoof, J., Valcke, M., & Van Petegem, P. (2010). Using school performance feedback: Perceptions of primary school principals. *School Effectiveness and School Improvement, 21*(2), 167–188. http://dx.doi.org/10.1080/09243450903396005.
Vescio, V., Ross, D., & Adams, A. (2008). A review of research on the impact of professional learning communities on teaching practice and student learning. *Teaching and Teacher Education, 24*, 80–91. http://dx.doi.org/10.1016/j.tate.2007.01.004.
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.
Weiss, C. H. (1998). *Evaluation: Methods for studying programs and policies*. Upper Saddle River, NJ: Prentice Hall.
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.
Yin, R. K. (2003). *Case study research: Design and methods*. Thousand Oaks, CA: Sage Publications.