



The effects of a data use intervention on educators' satisfaction and data literacy

Johanna Ebbeler¹ · Cindy L. Poortman¹ ·
Kim Schildkamp¹ · Jules M. Pieters¹

Received: 9 July 2015 / Accepted: 2 November 2016 / Published online: 14 November 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract Schools in many different countries are increasingly expected to use data for school improvement. However, schools struggle with the implementation of data use, because building human capacity around data use in education has not received enough attention. Educators urgently need to develop data literacy skills for being able to use data. For supporting schools with the endeavor of developing data literacy skills, we developed and implemented a data use intervention in secondary schools based in the Netherlands. This study therefore focuses on the effects of this intervention on educator satisfaction with the intervention and their data literacy skills and attitude toward data use. This study uses a quasi-experimental research design and employs a mixed-methods approach with a data use questionnaire filled in by data team schools ($N = 9$) and comparison schools ($N = 42$), a satisfaction questionnaire filled in by data team participants ($N = 55$), pre- and posttest knowledge tests filled in by data team participants ($N = 36$), and interview data ($N = 11$) from three case study schools. The results show that the participants were, for example, very satisfied with the support received during the intervention. Also, respondents developed new data literacy skills and showed a more positive attitude toward data use. The results show how teachers can be supported systematically in data use in their educational practice. In the

✉ Cindy L. Poortman
c.l.poortman@utwente.nl

Johanna Ebbeler
j.ebbeler@utwente.nl

Kim Schildkamp
k.schildkamp@utwente.nl

Jules M. Pieters
j.m.pieters@utwente.nl

¹ Faculty of Behavioural, Management and Social Sciences, ELAN: Department of Teacher Development, University of Twente, P.O. Box 217, 7500AE Enschede, The Netherlands

conclusions, we discuss some important implications for practice regarding the intensity and duration of support and implications for further research.

Keywords Data use · Professional development · School improvement · Mixed methods

1 Introduction and conceptual framework

Data use for educational decision-making has become prevalent in schools in many parts of the world (Cosner 2014; Datnow et al. 2013; Mandinach and Gummer 2013; Schildkamp et al. 2014). The information that is gained from data can be used to guide teaching, as well as learning processes (Halverson 2010). As a result, educators are increasingly expected to access and use data (Marsh and Farrell 2015; Piety 2013).

In the USA, the development toward data use in schools has started in the early 1990s and has resulted in the creation of large-scale information systems that collect, process, and store data (Anagnostopoulos et al. 2013; Piety 2013). Although schools have a multitude of qualitative and quantitative data readily available (e.g., observation data to represent quality of instruction in classrooms, student voice data to represent attitudes of students toward homework, or assessment data to represent student learning) coming from a broad variety of sources (for example, statewide student information systems, education agencies, or newspapers) (Anagnostopoulos & Bautista-Guerra 2013; Piety 2013), research about data use in schools shows that summative data, specifically student achievement data, are most commonly used by educators (Halverson and Thomas 2007; Shen et al. 2010). However, student achievement data alone provide only little information about the reasons behind the achievement results or about useful strategies that can support learning (Anderson et al. 2010; Halverson 2010). Therefore, we take a broad view on the term data and define data as “information that is systematically collected and organized to present some aspect of schools” (Schildkamp and Lai 2013). This includes not only achievement data but also, for example, classroom observations.

Although data are available in schools, and data use can lead to increased student achievement (Carlson et al. 2011; McNaughton et al. 2012), many decisions in schools are still based on intuition and limited observations (Ingram et al. 2004). Data literacy skills of educators are of critical importance if schools want educators to use data. However, often, educators lack the needed data literacy skills (Farley-Ripple and Cho 2014; Marsh 2012). Thus, building human capacity around data use in schools is necessary (Mandinach and Gummer 2013). To provide professional development (PD) in data use in secondary education, a data use intervention was developed in the form of data teams. This study focuses on the effects of participating in a data team on educators’ professional development regarding data literacy skills on a larger scale.

1.1 Data use theory of action

To further specify what kind of data literacy skills are needed by educators to use data in their schools, it is necessary to specify how schools can effectively use data. As presented in Fig. 1, data use involves an interpretative process, in which data have to be accessed, collected, and analyzed to be turned into information, and must be combined

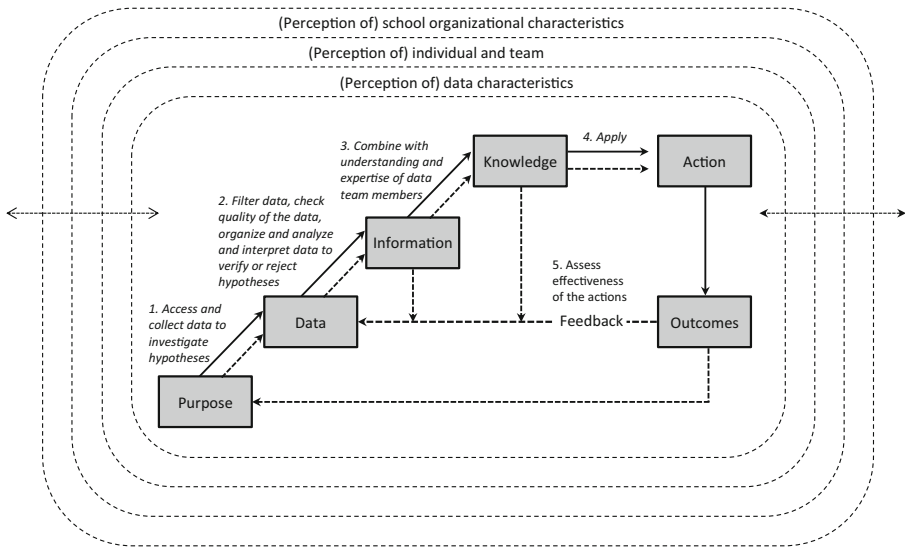


Fig. 1 Data use theory of action and factors influencing data use (Schildkamp & Poortman, 2015, based on Marsh 2012, p. 4; Coburn and Turner 2011; Ikemoto and Marsh 2007; Schildkamp & Lai, 2013; Mandinach et al. 2008; Schildkamp & Kuiper, 2010; Schildkamp & Lai, 2013; Supovitz 2010)

with understanding and expertise to become meaningful and useful for actions (Coburn and Turner 2011; Marsh 2012). Based on Marsh (2012, p. 4), and incorporating relevant characteristics of other data use models and frameworks (Coburn and Turner 2011; Ikemoto and Marsh 2007; Schildkamp and Lai (2013); Mandinach et al. 2008; Schildkamp and Kuiper 2010; Supovitz 2010), a data team theory of action (Fig. 1) was developed (Schildkamp and Poortman 2015). In this framework (Fig. 1), the interaction between data and people, in a certain context, results in decisions with regard to what action to take.

1.2 The data use intervention

For implementing data use in schools, the data use intervention in the form of data teams was designed. This data use intervention aims at enhancing educators’ data literacy skills about data use by giving educators a structured approach containing eight steps (see Fig. 2). This eight-step approach supports the team members to solve a problem emerging in their own school context by using (qualitative and quantitative) data, and it supports the active involvement of members (Bryk et al. 2015). Engaging educators in conversations regarding educational problems within their daily practice by using data creates powerful professional development opportunities, builds collegiality, and helps building professional relationships (Coburn and Turner 2011; Halverson 2010; Piety 2013; Brocato et al. 2014). The main objective of the data use intervention is professional development regarding data use by collaboratively solving a realistic problem defined and owned by the data team members.

In this study, data teams consisted of 4–6 teachers, 1–2 team leaders/school leadership team members, and if available in the school, also an internal data expert. These team members work together to solve an urgent educational problem in their own school context. Working according to the data team intervention gives educators within

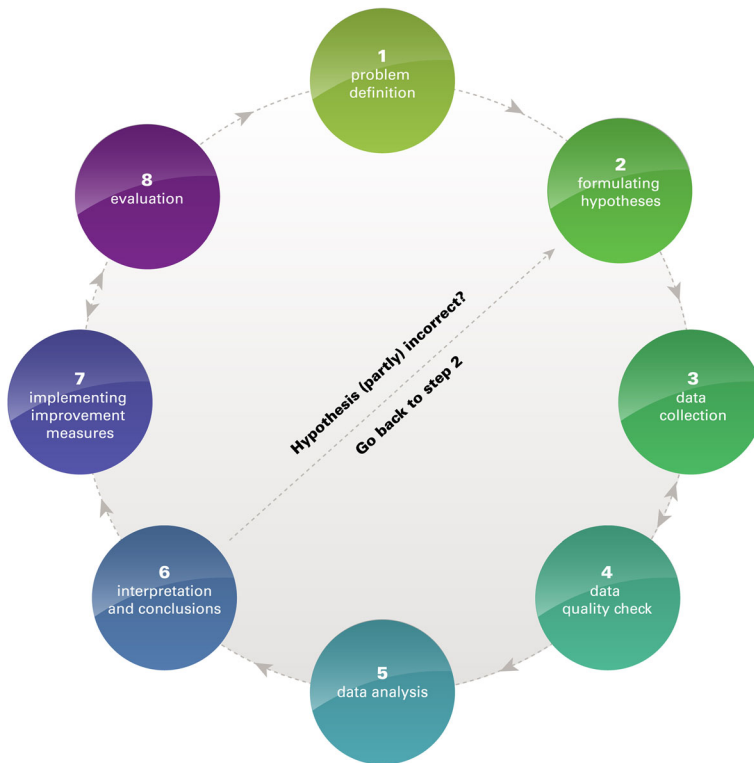


Fig. 2 The eight steps of the data use intervention (Schildkamp & Ehren, 2013, p. 56)

schools ownership about the process and makes educators active agents for data use (Bryk et al. 2015).

Data teams work according to a cyclic and iterative approach of the eight steps leading to the implementation of improvement measures based on data analysis. There is a general consensus about the steps that are important for effective data use in schools, albeit the steps vary across publications (see e.g., Boudett et al. 2005; Earl and Katz 2006; Marsh 2012). The eight steps in the data team intervention were inspired by existing data use manuals (e.g., Earl and Katz 2006; Boudett and Steele 2007). Furthermore, the intervention incorporated in the key elements of the data use theory of action (Fig. 1).

The eight step approach starts with a *purpose* by defining the problem (step 1). This means that the data team members brainstorm about educational problems within their school and goals they want to focus their efforts on. Making the problem specific is an essential step, because it gives a clear direction to data team members (Bryk et al. 2015). Next, data teams formulate concrete and measurable hypotheses about possible causes of the problem (step 2), for example, about possible causes for low examination results in mathematics. Subsequently, team members collect *data* to investigate the hypotheses (step 3). Qualitative and quantitative data collection methods can be employed for collecting, e.g., assessment data, examination results, or student voice data. In the next step, data team members check the quality of the data (step 4) regarding their reliability and validity. For example, members check to which degree

the instrument they used for data collection measured what it claims to measure. Next, members analyze the data (step 5) to be able to verify or reject the hypotheses under investigation. This may involve statistical procedures (e.g., descriptive analyses and *t* tests) or the analysis of qualitative data (e.g., coding and summarizing data). Subsequently, data team members interpret the analysis of the data and draw conclusions based on the analysis (step 6) implying data are to be transformed into *information*. Together with the participants' expertise and skills, this information is turned into *knowledge*. This step also implies that if the tested hypotheses turn out to be false, new hypotheses have to be formulated and tested. In that case, teams have to go back to step 2. Next, the data team takes *action* by implementing improvement measures (step 7). Finally, the team evaluates the implemented measures (step 8) by establishing the *outcomes*. Here, the team determines if the improvement measures were effective and if the goals were met. The eight steps are completed only if the goals have been accomplished.

The data teams are coached in following the data use approach by an external data coach for a period of one-and-a-half years. The external data coach joins each data team for 90 minutes every 3 to 4 weeks. It is the external data coach's task to monitor the process of the data team and give just-in-time support.

The data use intervention also includes two voluntary data analysis courses for data team members. These courses provide relevant entry statistical knowledge and skills. The basics of statistical analysis are taught in the first course (e.g., mean, standard deviation, and graphical representations of data), and more advanced data analysis is taught in the second course (e.g., *t* test, correlation, and chi-square test). Also, both courses contain information on qualitative data analysis and how to conduct the quantitative analyses by using Excel.

This intervention meets many criteria of effective professional development as found by researchers in the field of professional development. These include collaboration in a professional learning community, shared vision and goals, related to daily practice, active participation, leadership, structure, and support (e.g., Borko 2004; Desimone 2009; Guskey 2000; Jimerson and Wayman 2012; Lomos et al. 2011; Stoll et al. 2006; Vescio et al. 2008).

1.3 Educator satisfaction and data literacy

According to Guskey (2000), "professional development is defined as those processes and activities designed to enhance the professional knowledge, skills, and attitudes so that they might, in turn, improve the learning of students" (p. 16). For being able to use data effectively, educators need to develop their data use attitude and data literacy skills (Marsh 2012). Data literacy skills are the skills and knowledge that are needed to effectively use data in schools (Gummer and Mandinach 2015). Developing a positive attitude toward data use, or "buy-in/belief in data" (Schildkamp and Kuiper 2010), is necessary for the implementation of data use in schools (Datnow et al. 2007). Having or developing a positive attitude toward data use in this study means that participants recognize the added value of analyzing and using data and of having evidence for the claims they make regarding problems in their school.

Taking into account the steps that educators need to complete in order to effectively use data in their schools as presented in the data use theory of action (see Fig. 1), data

literacy skills are, for example, knowledge and skills about accessing, collecting, and analyzing data for investigating hypotheses.

A prerequisite for being able to acquire data literacy skills in an intervention is the satisfaction of the participants about the intervention (Desimone 2009; Guskey 2000), e.g., that participants are satisfied about the material that is provided, the extra courses regarding data analysis that are offered by the university, and the support as provided by the external data coach during the data team meetings. Testing satisfaction about the intervention is important, because the higher the satisfaction of participating educators with the intervention, the more likely they will obtain data literacy skills, as provided in the intervention, which they can use in practice in order to improve education (Nir and Bogler 2008).

Consequently, in this paper, we focus on the effects of the data team intervention on educator *satisfaction* and on educator learning outcomes (*data literacy skills and attitude*) with regard to data use. To evaluate the effects of participating in a data use intervention on educators' professional development regarding data use, our two research questions are as follows:

(RQ1) To what extent are educators *satisfied* with the data use intervention?

(RQ2) To what extent have educators' data literacy skills and attitudes improved after participating in the data use intervention?

2 Method

In order to be able to link results about educators' professional development regarding respondents' initial satisfaction (RQ1) and their data literacy skills and attitude (RQ2) to the data use intervention, a quasi-experimental research design was applied (Shadish et al. 2002) in which we used a mixed-methods approach as outlined in Table 1. The quasi-experimental research design was appropriate in this study, because data teams were implemented into the daily practice of the participating schools. Also, the participating schools were not randomly assigned to a condition, but applied for having a data team in school.

Table 1 Outline of the study

Group	Pre-measurements (January 2012)	Measurements during intervention	Post- measurements (June 2013)
Experimental group: data team schools (10 schools)	<ul style="list-style-type: none"> • Knowledge test • Data use questionnaire 	Data team participation: <ul style="list-style-type: none"> • Observation recordings • Data team evaluation with the external data coach 	<ul style="list-style-type: none"> • Knowledge test • Data use questionnaire • Satisfaction questionnaire • Interviews
Comparison group: schools without a data team (42 schools)	<ul style="list-style-type: none"> • Data use questionnaire 	(No participation)	<ul style="list-style-type: none"> • Data use questionnaire

2.1 Context

This study took place in the Netherlands. The Dutch Inspectorate holds schools accountable for the education they provide. Schools can use several data sources in order to improve their education, including, for example, school inspection data; school self-evaluation data; data on intakes, transfers, and dropouts; final examination results; assessment results; and student and parent questionnaire data (Schildkamp and Kuiper 2010).

The policy emphasis for data use has increased significantly in the Netherlands in recent years, as is the case internationally (Schildkamp et al. 2014). The Dutch School Inspectorate is increasingly holding schools accountable for using data to improve their quality (Verbeek and Odenthal 2014). It is the aim of the Ministry of Education that by 2018, at least 90% of primary and secondary education schools in the Netherlands apply data use. In 2010, only 20% of the Dutch secondary schools already applied data use (Dutch Inspectorate 2011).

2.2 Respondents

This project is part of a cooperation with one of the largest Dutch school boards in secondary education. Ten schools from this school board voluntarily signed up for participation with a data team. The 42 schools of the school board that did not participate in a data team formed our comparison group (see Table 1). These 42 schools are a naturally more comparable group for the intervention group than schools from other boards, because these schools operate under one board with the same vision and general policy regarding data use.

Furthermore, we analyzed whether the two groups were comparable by calculating chi-square tests regarding the two demographic variables: the gender of respondents and the subject area of respondents (i.e., languages, science and mathematics, or other subjects, for example, art, drama, PE, and sport) (Table 2) (Field 2013). Chi-square tests of independence were calculated comparing the frequency of the gender of respondents

Table 2 Chi-square test of independence cross tabulation for gender and subject area

	Gender			Subject area			
	Male Observed (O) <i>N</i> Expected (E) <i>N</i>	Female Observed (O) <i>N</i> Expected (E) <i>N</i>	Total Observed (O) <i>N</i> Expected (E) <i>N</i>	Languages Observed (O) <i>N</i> Expected (E) <i>N</i>	Science– Math Observed (O) <i>N</i> Expected (E) <i>N</i>	Other Observed (O) <i>N</i> Expected (E) <i>N</i>	Observed (O) <i>N</i> Expected (E) <i>N</i>
Data team schools	O: 139 _a E: 145.8	O: 138 _a E: 131.2	O: 277 E: 277	O: 74 _a E: 68.7	O: 87 _a E: 88.3	O: 116 _a E: 120	O: 277 E: 277
Comparison group	O: 262 _a E: 255.2	O: 223 _a E: 229.8	O: 485 E: 485	O: 115 _a E: 120.3	O: 156 _a E: 154.7	O: 214 _a E: 210	O: 485 E: 485
Total	O: 401 E: 401	O: 361 E: 361	O: 762 E: 762	O: 189 E: 189	O: 243 E: 243	O: 330 E: 330	O: 762 E: 762

Each subscript letter denotes a subset of gender categories whose column proportions do not differ significantly from each other at the 0.05 level

and the affiliation of respondents (affiliation in terms of working in a data team school or school belonging to the comparison group) and comparing the frequency of the subject area of respondents and the affiliation of respondents. A non-significant interaction was found for gender ($X^2(1) = 1.043$, $p = 0.307$) and subject area ($X^2(2) = 0.878$, $p = 0.645$). Therefore, we can conclude that the groups are comparable with regard to these aspects.

The problems these data teams were working on included, for example, the high retention rate in the fourth grade of senior secondary education (providing access to polytechnics) and the disappointing final examination results for English. Three data teams were selected for a qualitative case study during one-and-a-half school years, in which they were supported by the external data coach.

In order to select the three case study schools at the start of the project, first, a cluster analysis was run on all schools with among them nine schools having a data team, each responding to items about data use actions in school (i.e., data use for instruction, school development, and accountability; for more information, see the section on the data use questionnaire and Ebbeler et al. 2016). A hierarchical cluster analysis using Ward's method (Burns and Burns 2008; Saunders 1994) produced five clusters, between which the means of data use actions were significantly different. The clusters ranged from low to high means of reported data use in the schools.

Furthermore, the selection was based on teams being comparable regarding the presence of at least one school leader, an internal data expert, and at least three teachers during the meetings. This resulted in a selection of three case studies from three different clusters, ranging from low to high average scores per school for data use (see Table 3 for more information).

Table 3 Description of case study data teams

School	# of meetings	# of teachers	# of SLTs	# of internal data experts	Problem statement	# of respondents in the interview
A (scored highest in cluster analysis)	18	6	2	1	Declining number of students passing the final year of pre-university education	2 teachers 1 SLT 1 internal data expert
B (scored lowest in cluster analysis)	15	4	2	1	Disappointing final examination results of geography students	2 teachers 1 SLT 1 internal data expert
C (scored middle in cluster analysis)	16	6	1	1	Declining number of students passing senior secondary education (providing access to polytechnics)	2 teachers 1 SLT

2.3 Instruments

We used a combination of different instruments for triangulation purposes in order to adequately answer both research questions. In Table 4, we present all of the research instruments related to the research questions, as well as the data analysis conducted to answer the research questions.

2.3.1 Educator satisfaction questionnaire in the intervention group

For educator *satisfaction* (RQ1), we administered a short satisfaction questionnaire to the data team members ($N = 55$; 93.2% of all data team members) at the end of the support period. The evaluation questions were based on Guskey (2000), who recommended to ask questions with regard to satisfaction that are classified into categories, such as content questions related to utility and relevance of the activity, process questions related to organization of the activity, and context questions related to the setting of the activity. The questionnaire was distributed by the external data coach during one of the last data team meetings and returned to the researchers in a sealed envelope. This questionnaire included 21 items (see Table 5 for example items) regarding (1) support, (2) materials, (3) completing the steps in the intervention, and (4) the progress and process of data team meetings. All items were set on a five-point Likert scale ranging from “completely disagree” (1) to “completely agree” (5). Because not all teams had completed all steps at the end of the support period, we added the option “not applicable to all items” regarding the scale completing the steps in the intervention.

Table 4 Instruments and analysis

Research question	Respondents	Instrument(s)	Analysis
RQ1: To what extent are educators satisfied with the data use intervention?	All data team members	Educator satisfaction questionnaire	Descriptives (mean and se)
	Case study respondents	Data team evaluation with the external data coach	Qualitative data analysis
	Selection of case study respondents	Semi-structured interviews	Qualitative data analysis
RQ 2: To what extent have educators' data literacy skills and attitudes improved after participating in the data use intervention?	Pre-test and posttest for all data team members	Knowledge test	One-way
	between-subjects ANOVA Pre-test and posttest for intervention group and comparison schools	Data use questionnaire	Independent t test
Selection of case study respondents	Semi-structured interviews	Qualitative data analysis	

Table 5 Reliability of the scales in the satisfaction questionnaire

Scale	# of items	Cronbach's alpha	Example items
Support	4	0.85	The external data coach responded adequately to questions and concerns that were stated by the data team.
Material	3	0.85	I am satisfied about the content of the data team intervention manual.
Completing steps	8	0.74	We have completed step 1—problem definition—satisfactorily.
Progress and process of data team meetings	5	0.75	From my point of view, the data team has collaborated effectively.

Confirmative factor analyses and reliability analyses were carried out in SPSS, using principal component analysis and varimax rotation (Field 2013). The analysis revealed a four-factor structure: material (as provided during the training or data analysis courses); support (provided by the external data coach during meetings or the data analysis courses); completing steps in the intervention; and the progress and process of data team meetings. One item was deleted from the scale progress and process of data team meetings [2] because the item loaded lower than 0.5 (Field 2013). Reliability of all the scales was sufficient to good (see Table 5).

2.3.2 Data team evaluation with the external data coach in the case studies

In addition to the satisfaction questionnaire, the data team and the external data coach evaluated the satisfaction of the data team after the first year of training through questions about the support from the external data coach (e.g., What do you think about the support you receive?), the material (e.g., What do you think about the content of the data team manual?), the organization of the meetings (e.g., What could be improved with regard to the organization of the meetings?), and the process and progress of the data team meetings (e.g., What do you think about the collaboration during the meetings?) (RQ1). The evaluations were audio recorded in the three case studies. Due to the possible bias of the evaluation with the external data coach regarding socially desirable answers, this evaluation was triangulated with the results from the satisfaction questionnaire and the semi-structured interviews.

2.3.3 Semi-structured interviews in the case studies

At the end of the training period, semi-structured 1-h interviews with a selection of case study team members ($N = 11$) were conducted by a researcher to collect data about their satisfaction with the data team (RQ1), their attitudes toward data use, and data literacy skills acquired during the intervention (RQ2). Three to four members per case study data team were selected for the individual interviews: one school leader, the internal data expert (if available on the team), and two teachers. The external data coach supported in selecting respondents for the interviews. She was asked to indicate which two teachers in the case study data team were best able to articulate their view on working in a data team.

The questions in the interview schedule were based on the theoretical framework and were validated by an expert panel consisting of three researchers with teaching experience. Small adjustments were made, e.g., concerning the order of the questions and the formulation of some questions. Regarding satisfaction (RQ1), respondents were, for example, asked to rate their data team on a scale from 0 to 10 in the interviews and to further explain the reasoning behind the chosen rating. Participants were also asked what they thought about being a member of the team and were asked to express their opinion about the data use intervention. With regard to data literacy (RQ2), participants were asked what they had learned by participating in a data team.

2.3.4 Knowledge test in the intervention group

It is important to not only use perception data when investigating the effects of data use interventions (Marsh 2012). Therefore, for the question regarding educator *learning* outcomes (RQ2), we administered a knowledge pre-test and posttest ($N = 36$) for the data team members. The pre-test was taken during the second or third team meeting; the posttest during one of the last meetings.

The knowledge test was based on the data team guidelines for supporting the team members and included 12 open questions and tasks related to the content of the data team intervention (the maximum score was 22 points). Table 6 provides an overview of the open questions and tasks answered by the respondents. To avoid bias related to marking open questions, an extensive protocol for scoring the answers of respondents on the test was designed (Erkens 2011). The knowledge test was developed and discussed with colleagues, tested with external data coaches, and tested by a group of (five pilot study) data team participants. The knowledge test was administered only to the data team members.

2.3.5 Data use questionnaire for all respondents

In addition, we administered a data use questionnaire as a pre-test and posttest to study data literacy and attitude (RQ2), both to data team schools (only for teachers; pre-test: $N = 277$, 38.8% response rate; posttest: $N = 243$, 38.51% response rate) and schools in the comparison group (only for teachers; pre-test: $N = 485$, 20.7% response rate; posttest: $N = 788$, 35.53% response rate). This questionnaire was administered in combination with a larger research project and consisted of 61 items. It aimed at measuring data use at schools and was filled out by any teacher at a school regardless of whether the person was taking part in the data use intervention. Factor analysis and reliability analysis with this questionnaire were already conducted in a previous study (see Table 7) and showed that reliability of all of the scales was sufficient to good (Schildkamp et al. 2016).

For this specific study, however, only the scale regarding data literacy skills and attitudes was relevant. This scale consisted of eight items with a “good” reliability of $\alpha = 0.80$ (Field 2013). The items were set on a four-point Likert scale ranging from “completely disagree” (1) to “completely agree” (4). Due to the early stage of data use in the Netherlands, the alternative “don’t know” was also included.

Table 6 Overview of the open questions and tasks from the knowledge test

Task #	Related to step #	Task description
1	1—problem definition	Respondents have to name three requirements for formulating a good problem statement.
2	3—data collection	Respondents have to name two specific data sources for a given hypothesis related to a problem of low achievement.
3	2—formulating hypotheses	Respondents have to formulate a hypothesis that is related to a given problem.
4	4—data quality check	Respondents have to give two quality criteria for data, including a description of the criteria.
5	3—data collection	Respondents receive the task to describe a way for collecting data for testing a given hypothesis.
6	5—data analysis	Respondents receive the task to give a short analysis of pre- and posttest results showing the mean of students' satisfaction about teacher differentiation.
7	3—data collection	Respondents have to describe the difference between qualitative and quantitative research.
8	5—data analysis	Respondents have to describe a way in which interview results for a given example can be analyzed.
9	6—interpretation and conclusions	Based on a given case, respondents have to give a conclusion regarding the results of a questionnaire.
10	6—interpretation and conclusions	Respondents have to give a conclusion based on a given data set related to students repeating grades.
11	7—implementing improvement measures	Respondents have to formulate an improvement measure based on the conclusion they have provided in the previous task.
12	8—evaluation	Respondents have to describe how they can evaluate if the improvement measure as formulated in task 11 has worked.

2.4 Analysis

2.4.1 Educator satisfaction questionnaire in the intervention group

For the research question about educator *satisfaction* (RQ1), we used descriptive analyses to report the results of the satisfaction questionnaire administered to data team members. We analyzed the mean and standard error of each of the four scales (see Table 5) of the questionnaire.

Table 7 Reliability and example items of the scale of the data use questionnaire

Scale	# of items	Cronbach's alpha	Example items
Data literacy skills and attitude	8	0.80	<ul style="list-style-type: none"> • Students benefit when instruction is based on data. • I understand the quality criteria and concepts for data use (for example, correlation, validity, reliability). • Data are important in changing my teaching.

2.4.2 Data team evaluation with the external data coach and semi-structured interviews in case studies

The data team evaluations with the external data coach were audio recorded (as part of the observation recordings, see Tables 1 and 4) and transcribed verbatim. Also, the interviews with data team members of the case studies (for educator *satisfaction* and *data literacy skills, and attitude*, RQ1 and RQ2) were audio recorded. Interview summaries were sent to the individual respondents for a member check. All respondents agreed with the content. Next, the interviews were transcribed verbatim. The transcriptions of the evaluation with the external data coach and the interviews were coded by applying an a priori coding scheme (Strauss and Corbin 1998) based on the theoretical framework, using ATLAS.ti. Example codes include satisfaction with the process, attitude, and data literacy. Two researchers coded 10% of the same interview fragments. We calculated the inter-rater agreement and found an almost perfect Cohen's kappa of 0.83 (Landis and Koch 1977). After coding the interviews, a within-case analysis was conducted. Subsequently, a cross-case analysis was done.

2.4.3 Knowledge test in the intervention group

We conducted two types of analyses with the knowledge test (RQ2). To determine the reliability of this instrument, two researchers coded 21% of the same knowledge tests. The inter-rater agreement for grading the knowledge test was calculated using Cohen's kappa with linear weights and was found to be almost perfect at 0.92 (Landis and Koch 1977). A one-way between-subjects ANOVA was conducted to compare the effects of participating in a data team on the learning results between the data teams based on the gain scores resulting from pre- and posttest measurements. There was no significant difference regarding the effect of participating in a data team on educator learning between data teams at the $p < 0.05$ level for the gain scores of the knowledge test [$F(7, 28) = 0.966, p = 0.475$]. Thus, the ANOVA¹ indicated that it might not matter in which data team a respondent was participating, because the gain score did not significantly differ between the different data teams. Based on this outcome, all respondents of the data use intervention were treated as one group.

To assess whether the intervention helped to improve the data literacy skills of the respondents, we made use of the paired samples t test for assessing any differences between the pre-test and the posttest for data team members based on the gain scores of the knowledge test.^{2, 3}

¹ However, we do need to note here that it would have been better to take into account the nesting of respondents within schools by means of multilevel analyses. However, based on Hox (2010), we had too few respondents and schools to carry out multilevel analyses.

² After 6 months, one of the data teams chose not to proceed with the project because the intervention did not connect to the team leaders' goals. Their (pre-test) data regarding the knowledge test were therefore not included and no other data were collected there.

³ The item that was deleted from the scale *progress and process of data team meetings* in the satisfaction questionnaire was "I think that our data team did not reach as much as I hoped at the start of the intervention."

2.4.4 Data use questionnaire for all respondents

Also, concerning the research question about improved *data literacy skills and attitude* (RQ2), the data use questionnaire was analyzed, which was administered to teachers from both data team schools and comparison schools. Due to high overturn rates in schools and privacy reasons, it was not possible to match pre-test and posttest responses to the data use questionnaire at the individual level. Therefore, we have carried out independent *t* tests with the gain scores on the data use questionnaire variables for data team schools and comparison schools.

3 Results

In the results section, we will present the results per research question gathered by the different instruments (see Table 4). In the results section, we will use the term “educators” instead of “teachers” in the cases where the results apply to teachers, school leaders, and internal data experts together. The results regarding satisfaction will be presented in subsections according to the scales used in the educator satisfaction questionnaire (see Table 5). The results concerning the data literacy skills and attitude are subdivided into (1) *data literacy skills*, and (2) *attitude*.

3.1 Research question 1: effects on educator satisfaction

3.1.1 Support

The educator *satisfaction* questionnaire results [1] show that data team participants are, on average, satisfied to very satisfied about the *support* ($N = 55$; $M = 4.50$; $SD = 0.46$) during the data use intervention. This was confirmed by the results of the interviews. The support from the external data coach, as reported by interviewees from schools A and B, proved to be relevant, for example, because the external data coach made sure that the data team followed the method during the meetings (interviewees schools A and B), adequately structured the meetings (interviewee school A), and was a good team leader (interviewee school A): “[The external data coach] really made sure that we stick to the method” (interviewee from school A).

During the evaluation of the data teams with the external data coach, the team members in school C expressed that they valued the external data coach’s knowledge about the process of other data teams, which helped teams to make decisions about their own process.

Furthermore, interviewees in school B made statements about the frequency of the support as provided by the external data coach. An interviewee from school B reported: “...sometimes, the break between the meetings was too long, and I forgot what we were doing. I thought it would be a more continuous thing, and not like ‘Oh yes, we have another meeting. Let’s see, we had to prepare something...’” (interviewee from school B).

Here, we should note, however, that some of the meetings in this school were canceled, mostly on the school’s initiative, due to other priorities of the teachers. Data teams A and C reported during the evaluation with the external data coach that the

frequency of the meetings was good. In conclusion, the data teams were satisfied with the support they received from the external data coach.

Next, respondents from all schools reported their opinion about the data analysis course as provided by the university during the interviews. During the interviews in school A, one interviewee said that the course was good for refreshing prior knowledge regarding data analysis. Though one of the interviewees from school C appreciated the support from the university regarding the data analysis course, the respondent felt at the same time that this kind of support had little to no effect on the data team in that school: “Maybe, what I liked, (...), we followed the course [data analysis]...but in the end this did not affect our team meetings because we were the only two who had been there. I guess if the whole team had been there, it would have helped a lot...” (interviewee from school C).

3.2 Material

The satisfaction questionnaire results show that data team participants were, on average, satisfied to very satisfied about the materials ($N = 55$; $M = 4.14$; $SD = 0.62$) used in the data use intervention. The interviews with data team members showed similar results. In school A, for example, an interviewee stated that he appreciated the material as provided by the external data coach during the data team meetings, and also recognized that the material was updated during the intervention.

3.2.1 Completing the steps of the data use intervention

Respondents of the satisfaction questionnaire were neutral to satisfied about completing the steps ($N = 55$; $M = 3.88$; $SD = 0.49$). None of the three case study schools finished the data use intervention within the period in which they were supported by the data coach (i.e., one-and-a-half years). One interviewee in school C explained that data analysis is a time-consuming task. This may explain the moderate results of the satisfaction questionnaire regarding completing the steps. Also, in the evaluation with the data coach, some data team members said that because they did not complete all steps, it was hard to conclude whether they liked working according to the data use intervention.

3.2.2 Progress and process of the data team meetings

Respondents of the satisfaction questionnaire were neutral to satisfied about the process in their data team ($N = 55$; $M = 3.96$, $SD = 0.53$). Although some interviewees in schools A and B stated that overall the meetings were good (interviewees schools A and B) and that they were happy with the collaboration of the data team members (interviewees school A) and the enthusiasm and “fun” of the team while working on the data team (interviewees school A), there were also frustrations that may explain the more moderate results of the satisfaction questionnaire regarding progress and process. In schools A and B, for example, several interviewees reported that discussions sometimes were quite lengthy, meetings were sometimes unstructured (interviewees school B), and meetings could have been more efficient (interviewees school A). The data teams of schools A and B both describe a “dip” during their process: in school A,

there was a dip during the first phase of the process, but eventually it got better again; in school B, one interviewee described the final meetings of the data team as being slow and unstructured: "...but I have really got the feeling, that during the last meetings, that our pace is declining [during the meetings] and that we discuss anything and everything" (interviewee school B).

Despite the fact that some respondents experienced a lack of progress and were frustrated about not having completed the entire cycle, others thought the process as such was already a result, and also did not think that speed was a priority: "I think it is good that there was much space for discussion. Everyone could do his say and only then it is possible to get everything clear. This way we could make sure that everyone thinks the same. We wanted to do everything as thoroughly as possible. To me, it was fine that the discussion took so long" (interviewee school A).

The evaluation of the data team with the external data coach also underscores these questionnaire and interview results. Data team members of all case studies experienced the progress and process of the team (at times) as slow, also because meetings were canceled, team members were unprepared, or due to absence of members.

3.3 Research question 2: effects on data literacy skills and attitude

Inspection of the gain scores from pre-test to posttest survey results indicate that at the end of the intervention period, mean scores for the data literacy skills and attitude increased more for teachers in data team schools ($M = 0.10$; SE SD = 0.0412) than for teachers in the comparison group schools ($M = -0.06$, SE SD = 0.0526). The results of the independent samples t test show that these differences were statistically significant $t(40) = -1.747$, $p = 0.04$ (see Table 8). The effect size $d = 0.6$ represents a medium effect (Cohen 1992; Field 2013).

Regarding *data literacy skills* as measured by the knowledge test, data team members scored higher on the posttest ($M = 10.4$; SD = 3.05) than on the pre-test ($M = 9.4$, SD = 3.16). This difference, -1.04 , BCa 95% CI $[-1.82, -0.27]$, was significant $t(35) = -2.72$, $p = 0.005$. The effect size $d = 0.32$ represents a small to medium effect (Cohen 1992; Field 2013).

In the interviews, respondents from school A reported that the data team members have learned how to set up a good questionnaire by (re-)designing a

Table 8 Results of independent samples t tests and descriptive statistics for data use for data literacy skills, and attitude

Outcome	Group						95% CI for mean difference	t	df
	Data team schools			Comparison group					
	M^a	SD	n	M^a	SD	n			
Data literacy skills, and attitude	0.0971	0.1177	9	-0.0572	0.2559	33	-0.3328, 0.02422	-1.747	40

The items in the scales were answered using a Likert-type agreement scale, with 1 = *completely disagree* to 4 = *completely agree*, with the alternative answer *I don't know*

^a Mean of the gain score

questionnaire through exchanging feedback on that questionnaire. Team members also reported they learned what a scatterplot is. In school C, one of the respondents explained that they have learned how important data are to better formulate goals when trying to improve something in school.

Another respondent in school B gave a specific example of improving his data literacy skills: “I never worked with Excel for conducting statistical analyses. That is what I have learned: to calculate p values in Excel. And to be able to do it is really good” (interviewee school B). Also, data team members reported that they learned, e.g., how to calculate correlations in Excel (interviewee school B), that working with the data management system can be quite difficult (interviewee school B), and they learned about qualitative data analysis in the data analysis course (interviewee school C).

The interviews indicated that data team members’ data literacy skills have improved. However, the interviews also provide some explanation for why the effect size from the knowledge test was only small to medium. Several respondents stated that they did not learn anything really new with regard to data analysis (interviewees in schools A, B, and C). However, it is important to realize that respondents that gave this statement were mostly mathematicians, teachers with a background in sciences, quality care managers, or teachers who (recently) graduated from university with a master’s degree.

The interviews also gave an insight into the *attitude* of data team members regarding data use. For some respondents, like one in school B, using data for improving instruction was a new concept. Overall, interviewees of schools A, B, and C believed it was important to use data for decision-making and reported having a positive attitude toward data use. One of the interviewees in school B reported that it is good to have evidence before taking measures: “Maybe, we could have thought about these measures beforehand, but now we have evidence. That’s good. That’s ‘evidence based’” (interviewee of school B). This insight into the importance of using data increased during the intervention period. As one respondent stated: “I had no idea what it meant to use data, that was the reason for me to participate in the data team...and what I like about working in a data team is that you need to have proof. That you really need evidence. No more preconceptions, just facts” (interviewee of school C).

Also, an interviewee reported being more critical toward colleagues who are jumping to conclusions without having evidence. If colleagues gave solutions for a problem or assumed that they knew the cause for the problem, some data team members reported that they would ask these colleagues to prove these preconceptions (interviewee in school B).

However, there were also some critical remarks. In school B, one participant was not completely open to using data and did not fully accept the outcome of the data analysis. Also, in school A, there were some doubts about using data because the eight-step procedure had not been finished at the time of the interviews.

During the evaluation with the external data coach, data teams made some critical remarks as well, e.g., that the team will not continue working according to the eight-step approach (school A), and it became clear that not everyone on the team was convinced of the importance of using data for proving assumptions (school B).

4 Conclusions and discussion

The purpose of this article was to study the extent to which educators were satisfied regarding the data use intervention and the extent of improvement in educators' data literacy skills, and attitude regarding data use.

With respect to the research question about the extent of educators' *satisfaction* regarding the data use intervention, we can conclude that generally respondents were satisfied to very satisfied about the support in the data use intervention. Some aspects regarding the support could be improved, e.g., making schools more aware about the consequences of canceling meetings. Furthermore, data team members were moderately satisfied about their process, progress, and completing the steps of the intervention. Progress of the teams and completing the steps were hindered by both lack of momentum in the meetings and/or limited support time. The external data coach visited the schools for one-and-a-half years, which may have not been long enough. Studies show that professional development initiatives are more effective when they take place over a longer period of time (i.e., 2 years; Houtveen and Van de Grift 2012).

Regarding the second research question about the extent of an improvement in respondents' data literacy skills, and attitude regarding data use, we have found mixed to positive results. The educator learning results with regard to the *data literacy* knowledge test showed a small to medium effect, and with regard to the data use questionnaire, a medium effect. The effects on knowledge and skills might increase in the future by extending the training period and by improving the support and material (i.e., the guidelines) with more explicit principles and examples to help educators develop even more explicit data literacy skills. In the interviews, some accounts of data literacy skills development were given, for example, making scatterplots, conducting statistical analyses in Excel, and learning about qualitative research, but the qualitative data also show that not all participants felt they have learned something new with regard to data use in schools.

Regarding the attitude of respondents toward data use, we also found mixed to positive results. At some schools, data use was already (partly) implemented and some respondents stated that data use was part of their daily routine. Other respondents reported that participating in the data use intervention was their first encounter with data use and that they saw the benefits of using data. However, other educators were still not sure about the importance of using data at the end of the support period. In conclusion, we see that the attitude toward data use is changing to positive in most of the cases.

4.1 Implications for practice

This study shows that structured interventions such as this data use intervention can provide an important contribution to improving data use capacity building in schools. Crucial is that these data literacy skills are built using data from the teachers' own context and that teachers collaboratively use data. Unlike several other data use interventions (e.g., Carlson et al. 2011; McNaughton et al. 2012), this data use intervention starts with a problem the school chooses, and not with a general subject, such as literacy or mathematics.

Moreover, it is in the interaction within the team, between team members, and between the coach and the team members that based on data the improvement process starts. Through collaborative data discussions, taken for granted assumptions are disproven, myths are dispelled, actual causes of educational problems are found, and learning and improvement starts. This is what Earl and Timperley (2009) refer to as learning conversations. Some interventions use steps such as “digging into data” and “creating data overviews” (e.g., Boudett et al. 2005), and then move into the hypotheses (why) phase. The data teams move into the hypothesis phase sooner, to start with the learning conversations as soon as possible, to create a sense of ownership, and eagerness to learn.

Moreover, a structured step by step approach as used in this intervention is also important. The intervention has similar steps as the Plan Do Check Act cycle of Deming, and activities described in other data use projects and books (e.g., Boudett et al. 2005; Boudett and Steele 2007; Earl and Katz 2006). Data use with the goal of school improvement will always have similar components, including determining a purpose, data collection, from data to information to knowledge, to action, and evaluation (see also our data use theory of action). The results of this study show how important it is to make sure these steps are made explicit and concrete (and even then, it is still a challenge to develop data literacy).

Furthermore, the results of this study point to the importance of university–school partnerships. A coach from the university guided the team through all these steps. The data team coach supported the team, but the team also provided the data team coach and the university researchers with feedback. Based on the interaction with and feedback from the data teams, the data use intervention has been improved, for example, by adding more examples in the manual and by adding work sheets to the manual.

To summarize, important aspects for data use professional development interventions include the following:

- Teacher collaboration in a professional learning community. Similar to other interventions, teacher collaboration is the key to learning how to use data.
- Starting with a problem from practice, from teachers’ own context.
- Taking the hunches and ideas of participants seriously by collectively researching these in the form of hypotheses.
- Making all the data use steps as concrete and explicit as possible.
- University–school partnerships.

However, there are also challenges associated with these types of data use interventions. This study also shows that professional development requires time. Firstly, time for developing data literacy skills, and secondly, time for the required attitude shift. The attitude shift is needed, because educators need to feel the urge to use data. This shift cannot be created within a few meetings, but grows over time. Therefore, when developing data literacy, it is important to realize that these skills cannot be developed in just a couple of workshops, but that a long-term collaborative approach at the school site is much more likely to lead to the desired results.

The key challenge for practitioners is to share their data literacy skills with their other colleagues in the school and not only within a small team. The way in which

practitioners shape the process of knowledge sharing regarding data use will influence the sustainability of the intervention, and thus, also whether and how schools will use data for school improvement in the future. Therefore, we need further research into how data teams will use and share their new data literacy skills in practice.

Using and sharing data also requires a certain openness in the school and a level of trust. Only if educators are open and willing to use data to reflect on their own classroom practices, we can also expect improvements based on data within classrooms. This is especially important in low-performing schools, where educators often lack basic teaching skills (Mintrop and Sunderman 2013).

4.2 Limitations of the study and implications for further research

Some limitations of this study along with suggestions for further research should be noted. Future research should take into account that the knowledge test used in this study was only administrated within the intervention group and not in a comparison group. We tried to create a complete picture by triangulating the knowledge test results, with interviews and questionnaire results.

In this study, we showed that the participants have learned new data literacy skills. This study did not only use perception data in demonstrating these results, but we went beyond perception data by using a knowledge test to measure the effects on the knowledge and skills of participants. However, as other studies have shown, positive results at the learning level do not automatically lead to effects on student achievement (e.g., see Jacob et al. 2014). Future research should focus on the question of whether the participants actually *apply* their new data literacy skills in practice. Only if the new data literacy skills are applied in practice can improved student achievement, which is the ultimate goal of professional development, be reached (Kirkpatrick 1996; Guskey 2000). For example, some of the schools reported that their data team had managed to find the causes of their problem and started to implement measures to improve their student achievement. Further research is needed in order to assess whether these measures have had an impact on student achievement. From other studies, we know that this is possible, as these (quasi) experimental studies have found an effect of data use interventions on student achievement (e.g., Carlson et al. 2011; McNaughton et al. 2012).

In addition, as one of the respondents stated in an interview, the question is what happens after the external support ends. The aim is that teams are able to continue with data use after the support has ended. Therefore, not only short-term research into the effects but also research into long-term effects, i.e., the sustainability of data teams, is important.

Acknowledgements The authors would like to thank the Stichting Carmelcollege for making this project possible.

Compliance with ethical standards

Funding This project was funded by the Stichting Carmelcollege.

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Anagnostopoulos, D., & Bautista-Guerra, J. (2013). Trust and numbers: constructing and contesting statewide student performance information systems. In D. Anagnostopoulos, S. A. Rutledge, & R. Jacobson (Eds.), *The infrastructure of accountability* (pp. 41–56). Cambridge: Harvard Education Press.
- Anagnostopoulos, D., Rutledge, S. A., & Jacobson, R. (2013). Mapping the information infrastructure of accountability. In D. Anagnostopoulos, S. A. Rutledge, & R. Jacobson (Eds.), *The infrastructure of accountability* (pp. 1–20). Cambridge: Harvard Education Press.
- Anderson, S., Leithwood, K., & Strauss, T. (2010). Leading data use in school: organizational conditions and practices at the school and district levels. *Leadership and Policy in Schools*, 9(3), 292–327. doi:10.1080/15700761003731492.
- Borko, H. (2004). Professional development and teacher learning: mapping the terrain. *Educational Researcher*, 33(8), 3–15. doi:10.3102/0013189X033008003.
- Boudett, K. P., & Steele, J. L. (2007). *Data wise in action: stories of schools using data to improve teaching and learning*. Cambridge, MA: Harvard Education Press.
- Boudett, K. P., City, E. A., & Murnane, R. J. (2005). *Data wise: a step-by-step guide to using assessment results to improve teaching and learning*. Cambridge, MA: Harvard Education Press.
- Brocato, K., Willis, C., & Dechert, K. (2014). Longitudinal data use: ideas for district, building and classroom leaders. In A. J. Bowers, A. R. Shoho, & B. G. Barnett (Eds.), *Using Data in Schools to Inform Leadership and Decision Making (2014)* (pp. 97–120). Charlotte, NC: Information Age Publishing Inc.
- 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.
- Burns, R. P., & Burns, R. A. (2008). *Business research methods and statistics using SPSS*. SAGE Publications Ltd: London Extension chapters on advanced techniques accessed online: <http://studysites.uk.sagepub.com/burns/chapters.htm>.
- 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. doi:10.3102/0162373711412765.
- Coburn, C. E., & Turner, E. O. (2011). Research on data use: a framework and analysis. *Measurement: Interdisciplinary Research and Perspectives*, 9(4), 173–206. doi:10.1080/15366367.2011.626729.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159.
- Cosner, S. (2014). Strengthening collaborative data practices in schools: the need to cultivate development perspectives and diagnostic approaches. In A. J. Bowers, A. R. Shoho, & B. G. Barnett (Eds.), *Using Data in Schools to Inform Leadership and Decision Making (2014)* (pp. 67–93). Charlotte, NC: Information Age Publishing Inc.
- Datnow, A., Park, V., & Wohlstetter, P. (2007). *Achieving with data: how high-performing school systems use data to improve instruction for elementary students*. Los Angeles, CA: University of Southern California.
- 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. doi:10.1108/09578231311311500.
- Desimone, L. M. (2009). Improving impact studies of teacher's professional development: toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181–199. doi:10.3102/0013189X08331140.
- Dutch Inspectorate (2011). *The state of the art of education. Report about education 2009/2010*. [De staat van het onderwijs. Onderwijsverslag 2009/2010 (2011).] Utrecht: Inspectie van het onderwijs. www.onderwijsinspectie.nl
- Earl, L. M., & Katz, S. (2006). *Leading schools in a data-rich world. Harnessing data for school improvement*. Thousand Oaks, CA: Corwin Press.
- Earl, L. M., & Timperley, H. (2009). *Professional learning conversations: challenges in using evidence for improvement*. Dordrecht: Springer.

- 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. doi:10.1016/j.stueduc.2015.11.002.
- Erkens, T. (2011). Constructing open questions. [Het construeren van open vragen]. In P. Sanders (Ed.), *Examination at school [Toetsen op school]*. Stichting Cito Instituut voor Toetsontwikkeling: Arnhem.
- Farley-Ripple, E. N., & Cho, V. (2014). Depth of use: how district decision-makers did and did not engage with evidence. In A. J. Bowers, A. R. Shoho, & B. G. Barnett (Eds.), *Using Data in Schools to Inform Leadership and Decision Making (2014)* (pp. 229–252). Charlotte, NC: Information Age Publishing Inc.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. London: Sage Publications.
- Gummer, E. S., & Mandinach, E. B. (2015). Building a conceptual framework for data literacy. *Teachers College Record*, 117(4).
- Guskey, T. R. (2000). *Evaluating professional development*. Thousand Oaks, CA: Corwin Press.
- Halverson, R. (2010). School formative feedback systems. *Peabody Journal of Education*, 85(2), 130–146. doi:10.1080/01619561003685270.
- Halverson, R., & Thomas, C. (2007). *The roles and practices of student services staff as data-driven instructional leaders*. Madison, WI: Wisconsin Center for Educational Research.
- Houtveen, T., & Van de Grift, W. J. C. M. (2012). Improving reading achievements of struggling learners. *School Effectiveness and School Improvement*, 23(1), 71–93. doi:10.1080/09243453.2011.600534.
- 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, IL: National Society for the Study of Education.
- Ingram, D., Louis, K. S., & Schroeder, R. G. (2004). Accountability policies and teacher decision making: barriers to the use of data to improve practice. *Teachers College Record*, 106(6), 1258–1287. doi:10.1111/j.1467-9620.2004.00379.x.
- Jacob, R., Goddard, R., Kim, M., Miller, R., & Goddard, Y. (2014). Exploring the causal impact of the McREL Balanced Leadership Program on leadership, principal efficacy, instructional climate, educator turnover, and student achievement. *Educational Evaluation and Policy Analysis*, 0162373714549620.
- Jimerson, J. B., & Wayman, J. C. (2012). *Branding educational data use through professional learning: findings from a study in three districts. Paper presented at the American Educational Research Association Conference*. Canada: Vancouver.
- Kirkpatrick, D. (1996). Great ideas revisited. Techniques for evaluating training programs. *Revisiting Kirkpatrick's four-level model. Training & development*, 50(1), 54–59.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. doi:10.2307/2529310.
- 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. doi:10.1080/09243453.2010.550467.
- Mandinach, E. B., & Gummer, E. S. (2013). A systemic view of implementing data literacy in educator preparation. *Educational Researcher*, 42(1), 30–37. doi:10.3102/0013189X12459803.
- Mandinach, E., Honey, M., Light, D., & Brunner, C. (2008). A conceptual framework for data-driven decision-making. In E. Mandinach & M. Honey (Eds.), *Data-driven school improvement: linking data and learning* (pp. 13–31). New York, NY: Teachers College Press.
- Marsh, J. A. (2012). Interventions promoting educators' use of data: research insights and gaps. *Teachers College Record*, 114(11), 1–48.
- 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 Administration & Leadership*, 43(2), 269–289. doi:10.1177/1741143214537229.
- McNaughton, S., Lai, M., & Hsaio, S. (2012). Testing the effectiveness of an intervention model based on data use: a replication series across clusters of schools. *School Effectiveness and School Improvement*, 23(2), 203–228. doi:10.1080/09243453.2011.652126.
- Mintrop, H., & Sunderman, G. L. (2013). The paradoxes of data-driven school reform: learning from two generations of centralized accountability systems in the United States. In D. Anagnostopoulos, S. A. Rutledge, & R. Jacobson (Eds.), *The infrastructure of accountability* (pp. 23–40). Cambridge: Harvard Education Press.
- Nir, A. E., & Bogler, R. (2008). The antecedents of teacher satisfaction with professional development programs. *Teaching and Teacher Education*, 24(2), 377–386.
- Piety, P. J. (2013). *Assessing the educational data movement*. New York, NY: Teachers College Press.

- Saunders, J. (1994). Cluster analysis. *Journal of Marketing Management*, 10(1), 13–28. doi:10.1080/0267257X.1994.9964257.
- 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. doi:10.1016/j.tate.2009.06.007.
- 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, The Netherlands: Springer.
- Schildkamp, K., & Lai, M. K. (2013). Conclusions and a data use framework. In K. Schildkamp, M. K. Lai, & L. Earl (Eds.), *Data-based decision making in education: Challenges and opportunities* (pp. 177–192). Dordrecht, The Netherlands: Springer.
- Schildkamp, K., Karbautzki, L., & Vanhoof, J. (2014). Exploring data use practices around Europe: Identifying enablers and barriers. *Studies in Educational Evaluation*, 42, 15–24. doi:10.1016/j.stueduc.2013.10.007.
- Schildkamp, K., & Poortman, C. L. (2015). Factors influencing the functioning of data teams. *Teachers College Record*, 117(5). Advance online publication. <http://www.tcrecord.org/> Accessed 16 Feb 2016.
- Schildkamp, K., Poortman, C.L., Luyten, H., & Ebbeler, J. (2016). Factors promoting and hindering data-based decision making in schools. *School Effectiveness and School Improvement*. In press
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Wadsworth Cengage learning
- Shen, J., Cooley, V. E., Reeves, P., Burt, W. L., Ryan, L., Rainy, J. M., & Yuan, W. (2010). Using data for decision-making: perspectives from 16 principals in Michigan, USA. *International Review of Education*, 56(4), 435–456. doi:10.1007/s11159-010-9172-x.
- Stoll, L., Bolam, R., McMahon, A., Wallace, M., & Thomas, S. (2006). Professional learning communities: a review of the literature. *Journal of Educational Change*, 7, 221–258. doi:10.1007/s10833-006-0001-8.
- Strauss, A., & Corbin, J. (1998). *Basics of qualitative research techniques and procedures for developing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Supovitz, J. (2010). How principals and peers influence teaching and learning. *Educational Administration Quarterly*, 46(1), 31–56. doi:10.1177/1094670509353043.
- Verbeek, C., & Odenthal, L. (2014). Opbrengstgericht werken en onderzoeksmatig leiderschap in po en vo [data use and research leadership in primary and secondary education]. In M. Krüger (Ed.), *Leidinggeven aan onderzoekende scholen [leading researching schools]* (pp. 67–78). Bussum: Coutinho.
- 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(1), 80–91. doi:10.1016/j.tate.2007.01.004.