Distinguishing aspects of sustainability

A. P. M. Tappel[1] · C. L. Poortman[1] · K. Schildkamp[1] · A. J. Visscher[1]

Abstract

Many innovations that are implemented in schools are initially successful, but fail to become part of the schools' habits and routines. Relatively little research has followed innovations in schools for a long(er) time. In addition, few reforms last long enough to be studied longitudinally. In this exploratory study, the authors aim to find a way to quantitatively measure the degree of sustainability of a data use intervention three to eight years after its initial implementation. Sustainability in this study was defined as '…the process of integrating and scaling the innovation's core aspects in organizational routines that are adaptive to ongoing work, with continuing improvement of results'. A distinction was made between sustainability of the method and sustainability of the underlying goal of the data use intervention. A cluster analysis showed that schools can be categorized into four degrees of sustainability. A general view of the data indicated that the core components were present to varying degrees. Radar charts have been made and provided more insight. The process from organizational routines, and transfer of the intervention from the individual to the institutional level realizing organizational routines differs from school to school. The answer to the question as to whether an intervention has been sustained is not a simple yes or no. Sustainability is nuanced and can come in different forms. Measuring sustainability in a quantitative way contributes new perspectives and insights. With those insights further research can be done.

Keywords Sustainability · Data use · Data use for school improvement · Sustainable school improvement

✉ A. P. M. Tappel
a.p.m.tappel@utwente.nl

[1] Faculty Behavioural, Management and Social Sciences. Department of Teacher Development ELAN, University of Twente, Drienerlolaan 5, P.O. Box 217, 7552 NB Enschede, The Netherlands

Introduction

Many innovations that are implemented in schools are initially successful, but fail to become part of the schools' habits and routines (Wiltsey-Stirrnam et al., 2012). The toughest part of any educational innovation is not beginning it, but sustaining it in the organization (Hargreaves & Fink, 2012). The implementation of educational innovations takes a long time; however, relatively little research has followed innovations in schools for a long(er) time (Waslander, 2007). In addition, few reforms last long enough to be studied longitudinally (Datnow, 2005). Studies of educational sustainability have focused on continuity of school effectiveness based on statistical models (Creemers & Kyriakides, 2010), qualitative analysis of processes and innovation in schools (Giles & Hargreaves, 2006) and evaluations of programs and policies that seek to ensure the sustainability of results after completion of interventions (Datnow, 2005). Criteria for how schools can maintain these processes of change and improvement in student achievement over time are missing (Bellei et al., 2019) although there is a link with a strong and positive school culture (Lee & Louis, 2019). In-depth analyses of long-term innovation processes and of verified theories that explain and predict the long-term success or failure of an innovation are lacking (Cohen & Mehta, 2017; Mijs, 2007).

Data-based decision making is an example of educational innovation. Internationally, policymakers, researchers and practitioners increasingly recognize the significance of data use for school improvement (Datnow et al., 2013; Mandinach et al., 2015; Schildkamp & Kuiper, 2010; Visscher & Ehren, 2011). Data-based decision making involves collecting and using data to improve the quality of various educational processes and student learning and achievement (Van Geel et al., 2017). Studies have shown that the use of data can lead to improved student learning and achievement (Lai & McNaughton, 2016; Poortman & Schildkamp, 2016; Van Geel et al., 2016).

Despite the benefits of data use, it appears that many teachers do not use data that often, or use them incorrectly (Schildkamp & Lai, 2013). One reason is that teachers have difficulties with analyzing and interpreting data (Jimerson & Wayman, 2015; Kippers et al., 2018). Several data use professional development interventions have been developed to support teachers and school leaders in using data for school improvement, which target (a combination of) student learning, teacher learning, and organizational change (Coburn & Turner, 2012; Lai & McNaughton, 2016; Van Geel et al., 2016).

One such data use intervention is the data team intervention. This intervention has been introduced in over 60 schools in the Netherlands (Schildkamp et al., 2014), ranging from primary schools to institutions of higher education, but mainly in secondary education. In this intervention, teams of teachers and school leaders participate in an intensive professional development program for 1–2 years, to learn how to use data systematically. This data use intervention has two main goals: (1) the professional development of teachers and school leaders with regard to data use, in terms of the development of the relevant knowledge, skills and attitudes, and in terms of applying these competences in practice, (2) to improve educational quality.

Research has shown that this intervention can help improve educational quality (Ebbeler et al., 2016; Poortman & Schildkamp, 2016), but that its sustainability is a challenge (Hubers et al., 2017). Researchers have stressed that we need to understand what organizational structures and processes are required for the sustainable implementation of innovations and how these are experienced by educators (Wood, 2017). The innovation must become part of the daily organizational school routine to become sustainable. However, it appears that schools struggle to develop organizational routines for data use (Datnow & Park, 2015). The sustainability of the data use intervention seems to differ across the schools where it has been implemented, and needs further investigation. The purpose of this study is therefore to find out how sustainable this data use intervention is.

Theoretical framework

The data use intervention

School staff make educational decisions every day. When they make high quality, data-based decisions instead of decisions based on their intuition and experience alone, that can improve the quality of teaching and learning in the classroom (Schildkamp & Lai, 2013; Schildkamp et al., 2016). Making informed decisions is promoted by implementing data-based decision making, defined as making educational decisions based on different types of data (Lai & Schildkamp, 2013). Data are defined as 'information that is systematically collected and organized to represent some aspect of schooling' (Lai & Schildkamp, 2013, p. 10). Examples of data are assessment results, student satisfaction questionnaire results, or classroom observation results (Schildkamp & Lai, 2013).

The data use intervention discussed in this study is one way to support teachers and school leaders in effective data use (Andreou et al., 2014). It is based on several criteria for effective teacher professional development that have been frequently mentioned in the literature, such as collaboration between colleagues (van Veen et al., 2010), active leadership (Marsh, 2012a, 2012b), support by an expert (Marsh, 2012a, 2012b), and time to learn (van Veen et al., 2010).

In this intervention, teachers and school leaders work in what are called 'data teams'. A data team is a small-scale research team of 4–6 teachers, supplemented with 1–2 members of the school management. Supported by a trained coach, they together learn to work on solving educational problems at their own school, based on data. The data use intervention uses a structured, practical, step-by-step approach whereby activities are undertaken at every step, as described in a manual (Schildkamp & Poortman, 2015).

This step-by-step plan consists of:

1. *Problem definition* exploration of the problem the data team wants to focus on; thus, a topic that the school considers important and where school staff wants to see improvement. At this step, data are already being collected to determine how

   big the problem is. One goal for improvement in the short and longer term is also
   formulated.

2. *Hypothesis formulation* assumption about the causes of the problem the school is
   working on. Hypotheses are formulated as concretely and measurably as possible,
   in order to be able to substantiate whether the assumed causes identified in the
   hypotheses are actually the causes of the problem.

3. *Data collection* determine what data are needed to be able to test the hypotheses
   and where these data can be found, and collect the data. This can involve both
   qualitative and quantitative data.

4. *Data quality check* check whether the data found are of sufficient quality to be
   able to draw conclusions, by using the quality criteria of reliability and validity.

5. *Data analysis* thoroughly and carefully inspecting data, displaying data clearly
   and transparently and then summarizing them.

6. *Interpretation and conclusions* the data are interpreted in relation to the descrip-
   tion stated in the hypothesis, which is then accepted or rejected. In case of the
   rejection of the hypothesis, the data team goes back to step 2 to define a new
   hypothesis. If the hypothesis is accepted, the team can proceed to step 7.

7. *Implementation of improvement measures* the data team collects and chooses
   measures that meet criteria such as feasibility, proven effectiveness, speed of
   impact and costs. An action plan is drawn up, the method of evaluation is speci-
   fied and the measures are communicated and implemented.

8. *Evaluation* the measures are evaluated, both for the process, by monitoring the
   measures taken, and for the effect(s). Whether the measure has led to the solution
   of the problem is checked.

Data teams are supervised by a coach for either one year or two years. Data teams
coached for two years often focus on school-level problems, for example, grade rep-
etition by students (when students do not obtain the appropriate diploma within the
stipulated time). Most of the data teams that are coached for one year focus on a sub-
ject-specific problem, for example, disappointing exam results for a specific subject.
After the one or two year guidance from the coach, schools decide for themselves
whether and how to proceed with data teams and the use of data for school improve-
ment. Basic recommendations or guidelines regarding sustainability are provided
to the schools in the data team manual, for example starting spin-off data teams or
embedding the method in planned work.

   In the Netherlands, the data use intervention has been used in primary, second-
ary, and higher education since 2011 (after a 2-year pilot period). The data use
intervention has also been used in other countries, including Sweden (Schildkamp
et al., 2019) and the USA (Jimerson et al., 2020). The data teams in this study were
coached by an employee from the university where this method was developed.

Definition of sustainability

Research with a focus on sustainability of educational innovation has been con-
ducted within various forms of education: primary education (Larsen & Samdal,

2008; Payneeandy, 2012), secondary education (Furman Shaharabani & Tal, 2017; Kafyulilo et al., 2016), higher education (Elias, 2010; Roffe, 2010), and including several forms of education for children with special needs (Bambara et al., 2012; Benz et al., 2004). These studies focus on different units of analysis, ranging from the individual student level (Bambara et al., 2012), to program level (Ferguson et al., 2011; Furman Shaharabani & Tal, 2017), to the level of teachers (King, 2016) and their professional development (Edwards Groves & Rönnerman, 2013), school level (Andreou et al., 2014; Payneeandy, 2012), school district level (Alanís & Rodríguez, 2008; Sanders, 2012a) or school partnership level (Sanders, 2012b; Van Voorhis & Sheldon, 2004). However, most articles dealing with sustainability do not provide an explicit definition of the concept of 'sustainability' (Gaikhorst et al., 2018; Prenger et al., submitted). A number of dimensions have been mentioned in different articles however: sustained implementation of the *core components* of the intervention (e.g., Andreou et al., 2015a, 2015b; Drits-Esser et al., 2017; Gaikhorst et al., 2017; Kafyulilo et al., 2016); *continuation* of the intervention on the long run, or results that are maintained or continue to improve (e.g., Bean et al., 2015; Elder & Prochnow, 2016; Ferguson et al., 2011; Stringfield et al., 2008); becoming integral part of daily school *routines* (e.g., Bambara et al., 2012; Bean et al., 2015; Elder & Prochnow, 2016; Tam, 2009); *scaling up* of results (e.g., Dekker & Feijs, 2005); and *adaptiveness* (Benz et al., 2004; Deaney & Hennessy, 2007a, 2007b; Elias, 2010)*.* These elements will be discussed in order to define sustainability for this study.

The following definition of sustainability is used in this study:

> Sustainability refers to the process of integrating and scaling the innovation's core aspects in organizational routines that are adaptive to ongoing work.

The elements that form the definition will be discussed.

Fullan (1992) described sustainability as the final phase of a linear change process. In this final phase attention is paid to *continuation* of the innovation. Continuation is the first dimension of sustainability. In practice, however, change processes do not take place linearly in school organizations. Researchers have argued that sustainability is a developmental process that should start as soon as the initial stage of the change process (Fagen & Flay, 2009). The necessary conditions must already be in place during the adoption and implementation phases in order to guarantee sustainable innovation (Adams & Gaetane, 2011; Fagen & Flay, 2009; Pluye et al., 2005; Van den Boom-Muilenburg, 2021).

The second dimension of sustainability is enactment of an intervention must be visible in *the routines* within the organization (Bambara et al., 2012; Bean et al., 2015; Benz et al., 2004; Elias, 2010; Larsen & Samdal, 2008; Pluye et al., 2005). Implementation is only sustainable if people's new behavior is embedded in a daily routine (Waslander, 2007). Organizational routines are the main building blocks through which coordination, regularity and capabilities are generated in organizations (Dosi et al., 2000; Jacobs & Snijders, 2008) and 'recurrent and patterned interactions that guide engagement with data and people during ongoing work' (Coburn & Turner, 2012, p. 181)*,* which influence each other and require the involvement of multiple actors (Jacobs & Snijders, 2008). The innovation has to become part of the daily school routine (Bambara et al., 2012; Bean et al., 2015; Benz et al., 2004;

Elias, 2010; Larsen & Samdal, 2008; Pluye et al., 2005) in which new employees are also consciously included (Bambara et al., 2012).

Organizational routines can be considered as having *ostensive* and *performative* aspects (Feldman & Pentland, 2003a, 2003b) which both are important for sustainability. The ostensive aspect of an organizational routine is defined as the schematic, abstract idea of the routine (Feldman & Pentland, 2003a, 2003b), the'ideal' version of it. This involves standard procedures and established standards (Hubers et al., 2017). The performative aspect is defined as the specific actions that are undertaken to perform the organizational routine in practice (Feldman & Pentland, 2003a, 2003b) and refers to the everyday use of specific actions, carried out by specific people, at specific moments (Hubers, 2016). Organizational routines structure work practice and stabilize it over time. Change happens in the interplay between individual agency and the structure of the routine (Sherer & Spillane, 2011). The relationship between ostensive and performative aspects of routines creates opportunities for variation, selection, and retention of new practices and patterns of action within routines and allows routines to generate a wide range of outcomes, from stability to considerable change (Feldman & Pentland, 2003a, 2003b). New insights in organizational routines for sustainable educational innovation show a shifting focus from repetitive patterns of action, uniformity, and inertia, towards variations, adaptations, and change (Howard-Grenville et al., 2016).

Continuation of an intervention is an important dimension of sustainability, so are organizational routines. However, a critical eye for continuation of the intervention, which parts of the intervention will be retained or adapted, as well as on what scale the concept will be used remains important (Benz et al., 2004; Hargreaves & Goodson, 2006) to avoid meaningless procedures (Fullan, 2007).

A third dimension of sustainability is that working with the *core components* of the intervention happens during regular, ongoing work without interrupting existing practices (Coburn & Turner, 2012). The innovation is integrated with other initiatives in the organization (Hargreaves & Fink, 2000). Core components refer to components of the intervention itself or to the underlying, long term goal of the intervention. Core components of an intervention include the functions or principles and related activities necessary to achieve outcomes (Blase & Fixsen, 2013).

There is debate about the extent to which the concept guiding the original innovation must stay stable after implementation (often referred to as fidelity). In the literature, there is some tension between the flexibility and the stability of an intervention (Adams & Gaetane, 2011; Gaikhorst et al, 2018; Mitra, 2009). Sustainability is often automatically associated with the notion of stability. In traditional views, sustainable innovation is equated with the faithful implementation of the innovation (Mitra, 2009). The extent to which there is agreement between the original concept and the actual practice of the innovation is then seen as an indicator of sustainability. In this view, innovation must be introduced as faithfully as possible, with as few deviations as possible from the original plan in class and school practice (Mitra, 2009). So, the focus here is on 'continuing to do the same thing'. Recent research has criticized this traditional interpretation of sustainability (Adams & Gaetane, 2011), with a shift from stability to a more dynamic interpretation of the concept of sustainability (Fagen & Flay, 2009). Successful educational innovations are characterized in particular by the

fourth dimension, *adaptiveness:* with a view to successful and sustainable innovation, continual adjustments must be made with regard to the content and design of the concept guiding the innovation (Elias, 2010; Marsh, 2016). Innovations that are adapted to the specific context and needs of the school appear to be more successful in the long term with the continuation of the innovative program (Benz et al., 2004). Sustainable innovation is therefore about the extent to which the concept guiding the innovation is being developed further, and presupposes that there is an alignment and continual adjustment between the concept and actual class or school practice. Combining the two perspectives, sustainable innovation assumes a certain level of loyalty to the core components of the original innovation, but at the same time also incorporates the flexibility to adapt the concept guiding the innovation to (varying/changing) contexts (Gaikhorst et al., 2018). In short, the core components remain adhered to in practice (Sanders, 2012a), but, where necessary, can be adapted to the environment. What is needed is a balance between fidelity of implementation of the innovation and contextual adaptation (Sanders, 2012b).

Because long-time change in schools should not be restricted to only a few individuals, the use of an intervention (especially as manifested in its core components) should be noticeable within the school as a whole. Organizational routines also facilitate the fifth dimension of sustainability, *scaling up* of the results or the intervention. Scaling up has traditionally been viewed as increasing the number of teachers or classes using the intervention, or increasing the number of schools implementing the reform (Sanders, 2012a), which should take place by means of planned activities (Dekker & Feijs, 2005). Scaling up also assumes substantive change in educational practice resulting from the reform achieved by knowledge dissemination (Dekker & Feijs, 2005; Edwards Groves & Rönnerman, 2013). Knowledge dissemination is the transfer of knowledge to other areas, while scaling up can be seen as the transfer to larger groups of people. In order to be able to speak of sustainable innovation, we must therefore strive to maintain and disseminate high-quality change in teachers' teaching practices and a profound change in their views and norms (Adair Breault, 2013; Coburn et al., 2012). To support innovation and capacity it is necessary to find out under what circumstances, and how an innovation has been successfully sustained and scaled to new contexts (Howard et al., 2021).

Summarizing, we could state that sustainability is formed by four specific elements. First, the core components of the approach should be (second) continually carried out. Third the approach should become an integral part of the daily school routines, meaning that the approach is not perceived as something added or new, but as part of the "fabric" of the school. Fourth, the approach can be adapted over time while adhering to the core components, so it can be fitted to the work in the school. So sustainability is achieved when the core components of the professional development approach become a self- evident and functional part of the school (or: organizational routine), which is flexible and adaptive to ongoing work.

Sustainability of the data use intervention: Core components

The limited literature on sustainability in education mainly focuses on sustainable, often school-wide, implementation of (government-driven) way of working (Waslander, 2007; Prenger et al., submitted). There is relatively little longitudinal research, and the research available often follows pupils and rarely follows schools (Waslander, 2007). As the two main goals of the data use intervention are professional development by using the method and improving the school by the use of data, it is important to define the core components of the intervention to examine the sustainability of it in schools. The data use intervention is an iterative and cyclical eight-step procedure, with a supporting manual to be used in regular meetings. The eight steps, manual and regular meetings can be considered to be the core components of the intervention's *method*. The data use intervention is also an approach taken to achieve a larger goal in schools. Therefore, a distinction can be made between two dynamics of sustainability with regard to the intervention: the core components of the *method* and the core components of the underlying *goal* of the intervention. The core components of the method can also be defined as the specified components of the intervention.

The underlying goal of the data use intervention is the use of data by teachers and school leaders to improve the quality of their schools in the longer run. The core components of the underlying goal can be characterized as the general components of the intervention. Therefor they have to transpose the systematic way of working in the small data team to working with data for school improvement. The ability to understand and use data effectively to inform decisions is termed 'data literacy' (Mandinach & Gummer, 2013). Data literacy concerns educators' ability to set a purpose, collect, analyze, and interpret data, and take instructional action (Mandinach & Gummer, 2013; Schildkamp & Poortman, 2015; van Geel et al., 2017).

During the data use intervention's cyclical process, teachers and school leaders learn to use data systematically to improve the quality of education. Educators use these data literacy skills several times when following the eight steps of the data use intervention (Kippers et al., 2018). Five core components of the underlying goal of the data use intervention can be distinguished (Schildkamp & Poortman, 2015):

1. Setting clear, measurable goals;
2. Collecting data to determine whether the goals have been achieved;
3. And, if applicable, determining why goals have not been achieved;
4. Taking measures to achieve goals;
5. Evaluating the quality of education on the basis of data.

Sustainability of the data use intervention: Routines, continuation, stability versus adaptiveness of core components, and scaling up

Performative organizational routines with regard to the sustainability of the data use intervention refer to how people work with the data use intervention or its underlying goal in school teams. This means that there is a form of continuation

ROUTINE		SUSTAINABILITY OF THE METHOD	SUSTAINABILITY OF THE UNDERLYING GOAL
PERFORMATIVE	CONTINUATION	Still working with the method	Systematically using data for school improvement
	CORE COMPONENTS (stability *vs.* adaptiveness)	1. The eight steps 2. Use of the manual 3. Regular meetings	1. Setting clear, measurable goals 2. Collecting data to determine whether the goals have been achieved, 3. And, if applicable, determining why goals have not been achieved 4. Taking measures to achieve goals 5. Evaluating the quality of education on the basis of data
	SCALING UP	Quantitative increase	Qualitative change
OSTENSIVE		Individual job description	General job description – school plan

Fig. 1 Dimensions of the sustainability of the data use intervention

and scaling up of the (core components of the) intervention in practice. Ostensive organizational routines reflect the commitment expressed in policy, so whether or not procedures are defined and (job) responsibilities or tasks are documented in a formal way.

Regarding sustainability of the method, continuation means that schools are still working with the data use intervention. Regarding sustainability of the underlying goal, continuation means that schools are still systematically using data for school improvement, although no longer explicitly using the eight steps, manual or regular separate meetings to do so.

Regarding the sustainability of the core components of the method (eight steps, manual, meetings), stability refers to the extent to which these are used and adaptiveness refers to the extent to which the procedure has developed to fit into the specific school situation. As far as the core components of the underlying goal (setting measurable goals; collecting data; analyzing and interpreting; taking measures; evaluating on the basis of data) are concerned, adaptiveness refers to the extent to which these components are used in the school organization.

For the sustainability of the overall method, scaling up refers to any increase in the number of users of the data use intervention, for example, by starting new or spin-off data teams. For the sustainability of the underlying goal of the data use intervention, scaling up refers to the extent to which employees in the school systematically use data to improve its educational quality. Scaling up can then be seen as a qualitative change, involving not only the number of employees who are using data, but also the systematic use of data for quality improvement. Figure 1 summarizes dimensions of sustainability of the data use intervention.

In this study we aimed to answer the following research question: *How sustainable is the data use intervention in secondary schools in the Netherlands?* In answering this research question, we distinguished between the sustainability of the method and the sustainability of the underlying goal of the data use intervention.

Table 1 Overview of schools and respondents

# Schools/ respondents	School leader	Middle school leader	Teacher	Quality assurance employee
29	14	2	10	3
10	Not responding			
1	No data team members left			

Method

Context

We conducted this study in the context of Dutch secondary education. Dutch schools have the freedom to choose the principles (e.g., pedagogy, religion) on which the education they provide is based (Ministry of Internal Affairs, 2007). There is no national curriculum and teachers are free to develop assessments based on the curriculum determined by the school. National standardized assessments are taken only at the end of secondary education. Important data sources available within Dutch secondary schools and for the data use intervention include external data sources such as the results of those national standardized assessments and inspection data, as well as internal data sources such as curriculum-based assessments or student satisfaction questionnaires (Schildkamp & Kuiper, 2010).

Respondents

A total of 40 secondary schools that had been exposed to the data use intervention between 2012 and 2017 were approached to participate in this study. Their data teams had been funded by the Dutch Ministry of Education (17 schools), large school boards (19 schools), or by schools themselves (4 schools). Purposeful sampling was used for the identification and selection of respondents within these schools (Patton, 2002). This involves identifying and selecting individuals or groups of individuals that are especially knowledgeable about or experienced with a phenomenon of interest (Cresswell & Plano Clark, 2011; Palinkas et al., 2015). For this study, this means that the degree of sustainability of the data use intervention was investigated by consulting a specific expert within the schools, to obtain a clear picture of the sustainability of the data use intervention. The school leader involved with the data team was approached as the starting point for selecting the core respondents. In schools in which the school leader had left the school since the implementation of the intervention, the chairman or a different member of the original data team was approached. This could be an educational quality assurance employee (who has access to school data) or a teacher within the school. A core respondent could, for example, be a member of a spin-off data team or an employee with a focus on data use within the school. Table 1 gives an overview.

In the Netherlands, staff turnover in secondary education is about 17% of all teachers per year (Central Statistical Office). This influenced the absence of former

data team members in the schools that were invited to participate. A certain number of data team members and school leaders from the original data team had left the school, in all of the schools. Overall, almost 50% of the former data team members had left the school, for example, because of retirement, or switching jobs or schools. None of the former data team members still worked at one invited school. In addition, some schools did not respond to requests for cooperation or were not willing to participate for several reasons (10 schools). In most cases, schools were not willing to participate because the former school leader had left the school, or the school had already participated in other educational research. The number of schools within which research could be carried out was therefore reduced to 29. Respondents within these schools broke down by job category as follows: 14 secondary school leaders, 2 middle school leaders, 10 teachers, and 3 quality assurance employees.

Procedure and instrument

In this explorative study, we aimed to find a way to measure the sustainability of a data use intervention three to eight years after its initial implementation. Therefore we conducted a survey study (Floyd & Fowler, 2014) and collected data by means of a structured telephone survey in which the interviewer followed a set of questions in a predetermined order, with a limited number of response categories; the responses were recorded according a coding scheme (Denzin et al., 2013). Purpose sampling led to the choice for a telephone survey. The survey was based on the theoretical framework and included all aspects of sustainability, for both the *method* and the *underlying goal* of the method. These aspects are mentioned in the definition given and as elaborated in Fig. 1, namely, the performative organizational routine involving (1) continuation (stability vs. adaptiveness), (2) core components and (3) scaling-up of the intervention, as well as the ostensive organizational routine. The 14 questions were posed in a concrete way, to obtain a picture of the current situation in practice and to prevent socially desirable answers. For example, to examine the ostensive routine, respondents were asked whether the way of working used in the data use intervention was included in policy documents as an element of an employee's job description, or in the school plan. For continuation respondents were asked whether the school still works with the data use intervention (method/specified), and to what extent measurable goals are set to improve the quality of education (underlying goal/ general). Respondents were asked per core component of the underlying goals to what extent participation in the data use intervention has influenced this.

All respondents received the same questions in the same order, with the same answer categories. These were quantitative measurements. The answers were scored on a 4-point scale; a 'don't know option was added (see "Appendix" for the full set of questions).

To enhance construct validity (Cohen et al., 2013), the aspects were theoretically derived from reviews on sustainability (Gaikhorst et al., 2018; Prenger et al., submitted), among other things, and clarified in the theoretical framework of this study. Operationalization took place in collaboration with two researchers

who developed the data use intervention (Schildkamp et al., 2014). To avoid researcher bias, two researchers and nine teachers and school leaders from a school were consulted to check the formulation of the questions asked. Reliability for the survey instrument calculated by means of Cronbach's α was 0.97 (method) and 0.94 (underlying goal). To be able to study the connection between the data use intervention and the use of data for school improvement, respondents were explicitly asked whether participation in the data use intervention had influenced the use of data for school improvement.

Data analysis

To answer our research question, we conducted two types of analysis. First, to gain more insight into the data we had collected, we carried out descriptive analyses.

To gain insight into the *performative organizational routines*, we studied the continuation, stability and adaptiveness of the core components of both the method and the underlying goal, and scaling up.

- To gain insight into the number of schools continuing the data use intervention, we calculated the percentage of schools that were indicated to be still working with the data team intervention. Then, we investigated to what extent schools *scaled up* the intervention. For that purpose, we asked respondents how many data teams had been active in the past school year.
- In order to obtain insight into position regarding *stability vs adaptiveness* for both the *method* and the *underlying goal* of the method, means were calculated.

  - For stability-adaptiveness of the method, the score for the three core components (meetings, manual, eight steps) ranges from 1 (i.e., never) to 4 (i.e., always). A mean score of 4 on this newly constructed variable labeled stability-adaptiveness method means that the intervention has been implemented exactly as intended. We considered a mean score between 3 and 4 for the core components to indicate some degree of adaptiveness. That level of mean score indicates that data teams are still active, the eight steps and manual are being used on a regular basis and meetings take place at least every 8 weeks. A mean score lower than 3 was taken as indicating that the core components have not been implemented sustainably.
  - For the stability-adaptiveness of the goal of the intervention, the score for the five core components of the underlying goal of the intervention (set goals, collect, analyze, and interpret data, take measures and evaluate on the basis of data) ranges from 1 (i.e., nobody) to 4 [i.e., (almost) the entire school]. We considered a school to be adaptive if there were no active data teams, but the mean score for sustainability of the underlying goal was higher than 3. In that case, not just single individuals, but, for example, teams like subject teams, work with data for school improvement. We considered the way of working then as having been scaled up, because staff members are consciously working with data for school improvement.

- To measure the *ostensive routine*, we calculated the extent and presence of different types of policy documents in which the data team way of working with data for school improvement was captured. This is shown in percentages. A distinction was made between school plans (valid for the whole school or school location, meaning a 'this is how we work' routine), plans per team or subject department (applicable for specific teams, enactment), a personal development plan (as a personal goal to learn to use data for improvement of lessons), and in a job description (data use is expected to be done).

Second, a cluster analysis was conducted in order to group observations on the basis of similar characteristics in order to visualize the underlying structures in the data set (Hair et al., 1998); in this context, this was done to explore possible and meaningful clusters related to sustainability. First, a hierarchical cluster analysis using the Wards method was conducted in order to determine the number of clusters. We added all variables as summarized in Fig. 1; the performative organizational routine, which included continuation, the core components of both the method and the underlying goal and scaling up; and the ostensive organizational routine. The values for the ostensive routine and the core components of both the method and the underlying goals of the sustainability of the data use intervention were standardized. The dendrogram revealed four clusters. The second step in this procedure involved the determination of the sustainability clusters by conducting a K-means cluster analysis. To obtain more insight into the resulting clusters, the mean scores of each core component for both the method and the underlying goal were compared per category and plotted in radar charts.

Results

Sustainability of the method

Continuation and scaling up

In 14 of the schools studied (48%), the respondents indicated that school staff within their school location were still working with the data use intervention. Nine original data teams that had started in the 2012–2017 period were still active, and spin-off data teams had started in five schools. A spin-off data team is a newly composed team including one or more original members along with new members. One school started a new data team guided by an external coach (an educational consultancy organization), in order to continue the data team method of working within the school. All former data team members at this school location had left the school because of changing jobs or changing school locations, retirement or otherwise. The new school leader embraced the use of data for school improvement and reintroduced the method.

Summarizing, in terms of scaling up, of the 14 school locations that were still working with the data use intervention, five schools had two data teams that were

Table 2 Sustainability of the method: number of schools using the core components of the data use intervention by frequency of use, of the 14 schools still working with the intervention

	Core components of the method			
	Eight steps	Manual		Meetings
Frequency of use				
Never	1	1	Less than once a quarter	2
Sometimes	2	6	Once per quarter	7
Regularly	6	5	Every 8 weeks	3
Always	5	2	Monthly	2

still active, and nine schools had one active data team. There were no active data teams in the remaining 15 school locations.

Core components, stability versus adaptiveness

Among the 14 school locations that respondents reported to be still working with the intervention, one of the 14 schools had a mean score of 4 on the three core components of the method. In this school, the data use intervention had been implemented and was still used exactly according to the guidelines of the intervention. Five schools had a mean score of 3 or more, but less than 4, on the three core components (meetings, manual, eight steps) regarding the sustainability of the method. This means that the intervention was adapted to the school organization. In eight of the schools still working with data teams, the mean score was below 3. We consider mean scores below 3 as too far from the method to be considered to show sustainability.

A closer look at the general use of the core components (see Table 2) shows that using the eight steps had become a habit within the schools that still worked with the intervention and that this component was used in the most stable way of all three core components. The use of the manual varied, the great majority of schools (11) continuing to use the manual either sometimes or regularly, so it is getting used. Only one school had dropped it entirely.

Scheduling meetings within the school seemed to be a challenge; only two schools planned meetings as often as originally prescribed in the data use intervention.

Sustainability of the underlying goals

Continuation Fourteen of the 29 total schools participating had a mean score of between three and four on the core components of the underlying goal of the intervention (setting goals, collecting data, analyzing and interpreting data, taking measures and evaluating on the basis of data use). Even though respondents at 15 schools reported that they did not have active data teams at the time of the telephone interview, from nine of this 15 schools respondents indicated that data were used for school improvement by using these five core components, as a result of working with

data teams in the past. In the other five schools, the underlying goals of the data use intervention were not sustainable. So in total 24 out of 29 schools appeared to use data structurally for school improvement, even if they were no longer using the data use intervention as a method. The underlying goals of the data use intervention therefore can be considered sustainable in these schools. These schools have a routine for using data for school improvement in a flexible way.

Core components, stability versus adaptiveness Frequencies were calculated for all five core components related to the underlying goal (see Fig. 1) regarding whether they were used by nobody; individuals, such as former data team members; one or two teams; or (almost) the entire school or school location. Table 3 gives an overview of the frequencies and mean scores per core component of the underlying goal. Setting goals and collecting data seemed to be the most sustainable and stable core components. These components were most widely spread within the school. Analyzing why certain goals had not been reached, taking measures and evaluating educational quality on the basis of data seemed to be applied less widely.

Ostensive routine

Schools were asked to what extent their policy on data use was recorded on paper. Just over half of the 29 schools (15 schools; 52%) recorded the use of data for school improvement in their school plan or location plan. About half of the schools (12; 41%) included a data use policy in the plans of educational teams or subject departments. In about 15% of the schools, on average, data use was an element of job descriptions (5; 17%) or personal development plans (3; 10%).

Sustainability clusters

As shown in Tables 2 and 3, the use of the core components of the data use intervention differed within the schools studied and relationships between the core components were not visible. To determine whether the core components are interrelated, and to compare the use of these components in schools with different degrees of sustainability of the intervention, a cluster analysis was conducted. Schools were rated in terms of sustainability of the intervention. The cluster analysis resulted in four clusters of schools when it comes to sustainability of the intervention (see Table 4):

1. Cluster one can be labeled as 'intervention not sustainable' and includes seven school locations.
2. The second cluster can be called 'intervention sustainable on its method' and includes seven schools that still worked with the core components of the data use intervention, with the original or a new/spin-off team in which the method was used as intended. Still, only individual people worked towards the goal of the intervention (i.e., the data team members), and the ostensive routine was hardly visible in policy documents of the schools in this cluster.

Table 3 The sustainability of the core components of the underlying goal, N = 29

	Setting clear, measurable goals	Collecting data to determine whether the goals have been achieved	Analyzing and interpreting data	Taking measures to achieve goals	Evaluating the quality of education on the basis of data
Mean score per core components (n = 29)	3.0	3.0	2.7	2.7	2.7
Number of schools					
Nobody	1	2	5	5	6
Individually	8	8	9	8	7
1 or 2 teams/sections	9	8	5	6	6
Frequency of use					
(Almost) the entire school	11	11	10	10	10

Table 4 Four sustainability clusters

Cluster	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Classification	Intervention not sustainable	Intervention sustainable on its method	Intervention sustainable on its underlying goal	Intervention sustainable on both its method and its underlying goal
Continuation of the data use intervention	No	Yes	No	Yes
Scaling up	No	Original team—new/spin-off teams	Some individuals	Original team—subject or educational teams
Core components method	Negligible	Regular	Never—sometimes	Sometimes—regular
Core components goal	Some individuals	Individuals	1–2 educational or subject teams	1–2 educational teams up to the entire school
Ostensive routine	No	Hardly	No	Some
# Schools	7	7	9	6

3. Cluster three can be characterized as 'intervention sustainable on its underlying goal' and includes nine schools which did not work with the core components of the data team method anymore, but which worked cyclically (in parts of the school) on educational improvement using data.

4. Cluster four includes six schools in which (up to) the entire school location worked on the underlying goal of the intervention, and data teams were still active and following the method, and therefore can be called 'intervention sustainable on both its method and its underlying goal'. In this cluster, staff members did not use the core components of the method as strictly as described in the manual. The way of working was partially laid out in policy documents in these schools.

The separate frequencies as shown in Table 2 provide insight into sustainability of the method. To gain more in-depth insight, these frequencies were linked to the sustainability classification and made visible in radar charts, in which the mean scores for the core components of the method in the different clusters were compared and plotted (see Fig. 2).

(Cluster 1 is not visible; the intervention's method was not sustainable in these schools, and they therefore did not make use of the core components).

As shown in Table 2, the general use of the core components differed per component. However, by arranging the core components per cluster, more insight is gained into the sustainability of the core components of the method. As an example, in general, it can be said that the eight steps were used often in schools, but the plot shows that this was especially the case in schools in clusters two and four. School locations in cluster one (not sustainable) and cluster three (sustainable on the underlying goal), did not use, or hardly made use of the core components of the method. School locations in cluster two (sustainable on the method), make use of the three core components. School locations in cluster four also made use of the three core components, although the plot shows less frequent use of the manual and fewer meetings.
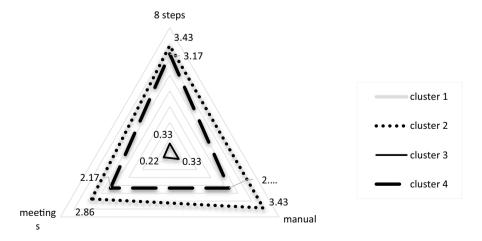

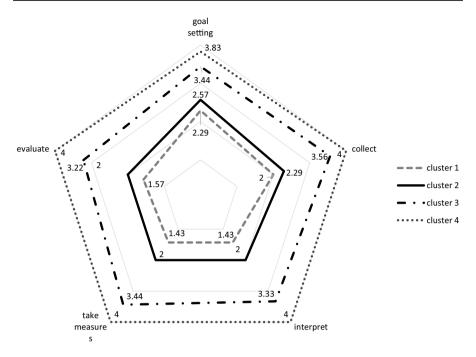**Fig. 2** Use of the core components of the method, per sustainability cluster

Fig. 3 Use of the core components of the underlying goal, per sustainability cluster

It seems that the manual is used if necessary in these school locations. The results related to the meetings core component show that scheduling meetings within the school seemed to be a challenge, especially in cluster four schools.

Table 3 shows how many schools followed the five core components of the underlying goal (goal setting, collect data, interpret data, take measure and evaluate on the basis of data), broken out by breadth of use within the school. To gain more in-depth insight into sustainability, these frequencies were linked to the sustainability classification and made visible in radar charts, in which the mean scores for the core components of the underlying goal in the different clusters were compared and plotted (see Fig. 3).

Schools in cluster four, where the intervention was sustainable on both its method and its underlying goal, scored high on the use of all five core components of the underlying goal, meaning that their use was relatively widespread within the school. School locations in cluster three (where the intervention was sustainable on its underlying goal) also used the five core components widespread, but slightly less than the schools in cluster four. Schools in these two clusters made working on school improvement on the basis of data an organizational routine. In cluster two schools, where the intervention was sustainable on its method, some individuals made use of the five core components of the underlying goal, but there was no organizational routine, given the low mean score. Cluster one schools, where the intervention was not sustainable, only some individuals made only little use of the core components. So, we could state that school locations in clusters

Table 5 Relation implementation year and cluster

Cluster/start	2012	2013	2014	2015	2016
1	2	3	0	1	0
2	3	1	1	1	1
3	5	4	0	2	2
4	0	2	0	0	1

three and four made data use for school improvement an organizational routine, in order to continually improve the quality of education at their school. School locations in clusters one and two had no organizational routine for working with data for school improvement. An important question here is also whether sustainability of data use in these schools was indeed related to the implementation of the data use intervention, or due to other initiatives within the school. We asked the respondents about this connection, and over 70% of the respondents indicated that their current data use practices were a result of participating in the data use intervention. Respondents indicated that with the implementation of the data use intervention they began to think about data, became familiar with the amount of (unused) data in the school, and started using data for school improvement.

The percentage of turnover of (key) school personnel did not appear to have a major influence between the clusters. The average leaving of data team members in cluster one was 41%, compared to cluster two 35%, cluster three 56% and in cluster four 37%. A closer look into clusters and starting year of the intervention shows schools that started with the data use intervention in 2012, 2013 and 2015 had a turnover about 44% and schools started in 2016 had a lower grade of 28%. These percentages show that, despite the high staff turnover, using data for school improvement in cluster 3 schools are indeed embedded in an organizational routine.

A closer look at the relationship between the year of implementation and the clusters did not provide a deeper insight either. See Table 5.

Conclusion and discussion

In this explorative study, we aimed to find a way to measure the sustainability of a data use intervention three to eight years after its initial implementation. Sustainability in this study was defined as *(…) the process of integrating and scaling the innovation's core aspects in organizational routines that are adaptive to ongoing work*. A distinction was made between sustainability of the method (specified components) and sustainability of the underlying goal (general components) of the data use intervention to get insight in sustainable use of the data use intervention itself as well as insight in the further development to use data for school improvement. Figure 1 gives an overview of the dimensions of sustainability for the data use intervention.

Profiles of sustainability

After conducting a cluster analysis on the results from 29 schools, four profiles for sustainability of the data use intervention were found.

1. *Intervention not sustainable*

    The cluster consists of seven schools (24%). The profile consists of the characteristics that define it. Seven schools could match this profile, but they could not constitute it. Schools in this profile neither continued to use nor scaled up the use of data teams. The use of the core components of the *method* was negligible and the use of the core components of the *underlying goal* only applied to a few individuals, usually former data team members. No policy on data use for school improvement was documented.

2. *Intervention sustainable on its method*

    Seven schools (24%) were aligned with this profile. Schools in this profile still had one or two active data teams, where the original data team was often still intact (continuation). Only a limited number of other school staff were making use of the data use intervention (i.e., limited scaling up). Original data team members often continued in a spin-off data team, with a couple of new members. The core components of the *method* (eight steps, manual and meetings) were used on a regular basis in these spin-off data teams. One school used the method exactly as intended. Other schools showed that they made a small adaption: the intervention prescribed meetings once every 3 to 4 weeks, the schools within profile 2 met on average once every 6 weeks. Data use in the profile 2 schools was not mandated in job descriptions or personal development plans very often, but in some schools, personnel was supported in using data for school improvement, so the ostensive routine was somewhat visible. Individuals in the schools in this profile made use of the five core components of the *underlying goal*, but this was not yet an organizational routine.

3. *Intervention sustainable on its underlying goal*

    Nine schools (31%) were aligned with this profile. Data teams were no longer active in the schools in this profile. Only some individuals in the school made use of the data use intervention; these are former data team members. The core components of the *method* were only sometimes used by a few individuals, but in these schools the cycle of setting a goal, collecting and interpreting data, taking measures and evaluating on the basis of data (*underlying goal of the intervention*) was a routine within several educational or subject-level teams. However, this was not an ostensive routine. Working with the core components of the underlying goal of the intervention was not documented in policy papers.

4. *Intervention sustainable on both its method and its underlying goal*

    Six schools (21%) were aligned with this profile. In these schools, data teams were active, but related work was also carried out within educational and/or subject-level teams. In terms of scaling up, more school personnel made use of the data use intervention. Besides the original data team, educational and/or subject-level teams also worked with the intervention. Schools in this profile made use of the three core components of the *method*; however, the manual seemed to be used

less often. An explanation for this could be that data team members trust their experience with working with the intervention, whether this is justified or not. There also seemed to be fewer official data team meetings. An explanation could be working according to the data use intervention was a routine in the school and the data team meetings were integrated in meeting cycles within the school, such as educational team meetings, subject meetings, consultation between class mentors and work groups. Therefore, meetings might not always be recognized as data team meetings. Exploratory interviews revealed this to be the case. The core components of the *underlying goal* (set goals, collect and analyze data, take measures, evaluate on the basis of data) were visible in profile four schools. These schools used these core components more widely; the organizational routine thus seemed to be more embedded. The ostensive routine was partially laid out in policy documents of these schools. The mandate of the use of data for school improvement was visible in a certain number of schools; half of the researched schools included a data use policy in their school plan or in plans for educational teams and subject departments.

In summary, the results of this study show that in 7 (24%) out of 29 schools included in the study, the data use intervention was not sustainable and some form of sustainability was found in 22 (76%) schools. Schools did not automatically continue, formalize or scale up this intervention in the same way or with the same goals. The results of this study clearly show that the answer to the question as to whether an intervention has been sustained is not a simple yes or no. It is important to adopt a more nuanced view when describing sustainability, and important to realize that sustainability can come in different forms.

This research also shows that sustainability of the data use intervention, even with quite a bit of external support, is not simple. The process from intervention to realizing organizational routines has proven to be difficult. On the basis of the number of school staff members working with the data use intervention, we could cautiously conclude that the transfer of the intervention from the individual to the institutional level differs from school to school and often proves to be challenging, as others have found as well (Akkerman & Bakker, 2011; Collinson & Cook, 2006; Gaikhorst et al., 2018).

Implications for practice

This study shows that sustainability involves more than the fidelity of implementation of an intervention, and also needs to include the underlying goal of that intervention. By making a distinction between the method and the goal, but also by distinguishing and describing these profiles of sustainability, the concept of sustainability becomes more nuanced and one becomes more aware of its various aspects. Even though in some schools the specific intervention itself wasn't applied anymore, its underlying goal was still maintained in terms of systematically using data. Schools reported this had been influenced by the original data use intervention. It is important to monitor the intervention, for example, with a checklist. A checklist can

consist of points of attention (e.g., goal, knowledge sharing) before, during and after implementation of the data use intervention. Monitoring and communication about these developments can stimulate the discussion about the added value of the intervention within the school organization (Hargreaves & Goodson, 2006).

In three out of four sustainability profiles, the data use intervention steps of interpreting data, taking action based on data and evaluating are used less often than the steps of setting goals and collecting data. It seems that schools work with data, but do not always use data in a beneficial way, and have difficulties with the last three steps in the cycle, as other data use studies have also demonstrated (Datnow & Hubbard, 2016; Marsh, 2012a, 2012b). Paying extra attention and possibly extra training in these steps might be necessary here. In addition, previous research illustrated that data team members gain knowledge about using data for school improvement (Ebbeler et al., 2016), but they struggle with brokering their knowledge to their colleagues (Hubers et al., 2018, 2019). Some studies have shown how teachers can face boundaries within their own team (Venkat & Adler, 2008) and when working with other teacher teams (Cobb et al., 2003). In order to realize sustainability, attention needs to be paid to knowledge brokerage and boundary crossing.

The high staff turnover in schools did not appear to negatively influence the embedding of data use for school improvement in organizational routines. Explanations can be found in sharing knowledge with new teachers (Andreou et al., 2015a, 2015b; Saito et al., 2012), collegial formal and informal personal contacts (Dekker & Feijs, 2005) as well as communication among staff (Elder & Prochnow, 2016; Kirtman, 2002). School leaders play a role (van den Boom-Muilenburg, 2021), so explicit attention of school leaders in planning, organizing and providing a clear direction form the basis of involving new school personnel in ongoing implementation of interventions in schools (Homan, 2017).

Recommendations for further research

This research has led to an elaboration and empirical measurement of the concept of sustainability in the context of educational innovations, which provides a basis for future research in this field. A limitation of this study is that the outcomes were based on questioning only one key person in the school, because several (original) data team members were not available in the school anymore. In order to be able to compare schools in a similar way, it was decided to administer the questionnaire to only one person at the school, even though the questionnaire is then based on self-report. To avoid social desirability, the questions were posed as concrete as possible, but this still is subjective. In order to obtain more complete insight into sustainability in schools working with this intervention, further research could focus on actually observing sustainability in the schools.

In this study we found different sustainability profiles; we see reason for further research in the variety of these profiles. It is unclear what factors influenced the process of sustainability. The literature points to many factors that could affect sustainability, such as teacher buy-in, leadership or effectiveness of the intervention (Gaikhorst et al., 2018; Prenger et al., submitted). To determine which factors or groups of

factors have the most influence on sustainability, more insight into these factors and their mutual influences is necessary (Wiltsey-Stirman et al., 2012).

Further research is also needed into the concept of organizational routines and what these routines look like within the organization. Organizational routines can be approached as single routines (Kremser & Schreyögg, 2016), but complementary routines form a system in which each routine contributes a partial result to the accomplishment of a common task (Kremser & Schreyögg, 2016). By focusing on the pedagogic-curricular elements in practice within a school, as this is the decisive factor in school performance, it can be determined at that layer to what extent an organizational routine is actually a routine (Bellei et al., 2019).

In most schools, the original data team was still working on educational problems. This implies that the knowledge about data use remains with a limited group of people. Research shows that not just knowledge, expertise and skills are key, but how these circulate in schools (Coburn et al., 2012; Daly & Finnigan, 2010) is crucial as well. Schools in profile four were found to be able to make a transfer from data team members individual knowledge to the institutional level. According to Daly and Finnigan (2010):

> Social capital is an investment in the social relations in a system through which the resources of other individuals can be accessed, borrowed, or leveraged. This differentiates social capital from human capital, which refers to investments in training, development, or certifications of individuals, or physical capital that is contained in infrastructure and equipment. (p. 115)

Therefore, it might be worthwhile to further investigate social networks within schools, for example, to study the differences in social networks between schools in profile four and the other profiles. These social networks can also help us by providing more knowledge about how to scale up an innovation from a small group of people to more or all school personnel. In order to sustain or scale up interventions in schools, it seems necessary to consciously plan and implement activities (Dekker & Feijs, 2005) and to start to think about how to sustain an intervention even before its implementation (Hubers et al., 2018).

Sustainability can be viewed in different ways. Most studies on sustainability have focused on fidelity, impact or effectiveness of the intervention for student achievement (Alanís & Rodríguez, 2008; Bean et al., 2015; Edwards Groves & Rönnerman, 2013; Elias, 2010, van den Boom-Muilenburg et al., 2019; Wolthuis et al., 2020), teacher change (Drits-Esser et al., 2017; Furman Shaharabani & Tal, 2017; Gaikhorst et al., 2017) or spreading knowledge of the intervention (Deaney & Hennessy, 2007a, 2007b; Dekker & Feijs, 2005). This study further examined what sustainability looks like by distinguishing the aspects of the process of sustainability. An important scientific contribution of this study is that we have started to pick apart the aspects of sustainability as a first step to understand its process. Variability between schools and context of schools makes it difficult to draw general conclusions in education (Mcnaughton, 2021). Sustainability also turned out not to be simply a matter of black and white, but is more nuanced and takes several forms. The profiles could help schools indicating if substantive change in educational practice resulting from the reform (Sanders, 2012a) has taken place.

Appendix: Survey questions, mapped in dimensions of sustainability, response options, and scoring system for sustainability

Survey questions

Name / function / school							
Year of introduction							
Organizational routine	Aspects of sustainability	Sub-aspects	Question	Response options (& points)			
Performative routine Core components Content intervention	Continuation		Do you still work with the data use intervention?	Yes (2) / No (1)			
	Scaling up	Quantitative	How many data teams were active last year?				
			Who is working with the data team intervention?	Original data team (1)	New or spin-off data team (2)	Subject / educational teams (3)	This school (location) (4)
	Stability Adaptiveness	Eight steps	To what extent do you still work according to all eight steps of the data team intervention?	Never (1)	Sometimes (2)	Regularly (3)	Always
		Manual	To what extent is the manual used?				
		Meetings	How often does the data team meet?	Less often than once a quarter (1)	Once a quarter (2)	Every 8 weeks (3)	Monthly (4)
Core components Underlying goal of the intervention: to achieve data use for educational improvement through a professional development intervention so that school development takes place	Continuation	Set measurable goals	To what extent are measurable goals set to improve the quality of education?	No one (1)	Individual (2)	1 or 2 subjects / teams (3)	(Almost) Entire school (4)
			Participation in the data use intervention has influenced this.	Totally disagree	Disagree	Agree	Totally agree
		Collect data	To what extent are data collected in order to determine whether the set goals have been achieved or not?	No one (1)	Individual (2)	1 or 2 subjects / teams (3)	(Almost) Entire school (4)
			Participation in the data use intervention has influenced this.	Totally disagree	Disagree	Agree	Totally agree
		Analyze and interpret data	To what extent are data used to determine why goals have / have not been achieved?	No one (1)	Individual (2)	1 or 2 subjects / teams (3)	(Almost) Entire school (4)
			Participation in the data use intervention has influenced this.	Totally disagree	Disagree	Agree	Totally agree
		Take measures	To what extent are measures based on these data being taken to achieve goals?	No one (1)	Individual (2)	1 or 2 subjects / teams (3)	(Almost) Entire school (4)
			Participation in the data use intervention has influenced this	Totally disagree	Disagree	Agree	Totally agree
		Evaluate	To what extent are data used to evaluate whether the goal has been achieved?	No one (1)	Individual (2)	1 or 2 subjects / teams (3)	(Almost) Entire school (4)
			Participation in the data use intervention has influenced this	Totally disagree	Disagree	Agree	Totally agree
Ostensive routine			Is this way of working included in the following policy documents?		Yes (2)		No (1)
				Employee job description			
				Personal development plan			
				Subject or educational team plan			
				School / location plan			

References

Adair Breault, D. (2013). The challenges of scaling-Up and sustaining professional development school partnerships. *Teaching and Teacher Education, 36*, 92–100. https://doi.org/10.1016/j.tate.2013.07.007

Adams, C. M., & Jean-Marie, G. (2011). A diffusion approach to study leadership reform. *Journal of Educational Administration, 49*(4), 354–377. https://doi.org/10.1108/09578231111146452

Adams, C. M., & Gaetane, J. M. (n.d.). A diffusion approach to study leadership reform. *Journal of Educational Administration, 49*(4).

Adams, C. M., & Jean-Marie, G. (n.d.). Journal of Educational Administration A diffusion approach to study leadership reform. *Journal of Educational Administration Journal of Educational Administration*. https://doi.org/10.1108/09578231111146452

Akkerman, S. F., & Bakker, A. (2011). Boundary crossing and boundary objects. *Review of Educational Research, 81*(2), 132–169. https://doi.org/10.3102/0034654311404435

Alanís, I., & Rodríguez, M. A. (2008). Sustaining a dual language immersion program: Features of success. *Journal of Latinos and Education, 7*(4), 305–319. https://doi.org/10.1080/15348430802143378

Andreou, T. E., McIntosh, K., Ross, S. W., & Kahn, J. D. (2015a). *Critical incidents in the sustainability of school-wide positive behavioral interventions and supports*, 157–167.

Andreou, T. E., McIntosh, K., Ross, S. W., & Kahn, J. D. (2014). Critical Incidents in sustaining school-wide positive behavioral interventions and supports. *The Journal of Special Education*. https://doi.org/10.1177/0022466914554298

Andreou, T. E., McIntosh, K., Ross, S. W., & Kahn, J. D. (2015b). Critical incidents in sustaining school-wide positive behavioral interventions and supports. *The Journal of Special Education, 49*(3), 157–167. https://doi.org/10.1177/0022466914554298

Bambara, L. M., Goh, A., Kern, L., & Caskie, G. (2012). Perceived barriers and enablers to implementing individualized positive behavior interventions and supports in school settings. *Journal of Positive Behavior Interventions, 14*(4), 228–240. https://doi.org/10.1177/1098300712437219

Bean, R. M., Dole, J. A., Nelson, K. L., Belcastro, E. G., & Zigmond, N. (2015). The sustainability of a national reading reform initiative in two states. *Reading and Writing Quarterly, 31*(1), 30–55. https://doi.org/10.1080/10573569.2013.857947

Bellei, C., Morawietz, L., Valenzuela, J. P., & Vanni, X. (2019). *School effectiveness and school improvement an international journal of research, policy and practice effective schools 10 years on: Factors and processes enabling the sustainability of school effectiveness*. https://doi.org/10.1080/09243453.2019.1652191

Benz, M. R., Lindstrom, L., Unruh, D., & Waintrup, M. (2004). Sustaining secondary transition programs in local schools. *Remedial and Special Education, 25*(1), 39–50.

Blase, K., & Fixsen, D. L. (2013). Core intervention components: identifying and operationalizing. *ASPE Research Brief, February*, 21.

van den Boom-Muilenburg, S. N., Daly, A. J., Poortman, C. L., Schildkamp, K., de Vries, S., & van Veen, K. (2019). *Sustainable school improvement with within-school PLNs: Exploring the position of leaders in the reform network*.

Cobb, P., McClain, K., de Silva Lamberg, T., & Dean, C. (2003). Situating teachers' instructional practices in the institutional setting of the school and district. *Educational Researcher, 32*(6), 13–24.

Coburn, C. E., Russel, J. L., Kaufman, J. H., & Stein, M. K. (2012). Supporting sustainability: Teachers' advice networks and ambitious instructional reform. *American Journal of Education, 119*(1), 137–182. https://doi.org/10.1086/667699

Coburn, C. E., & Turner, E. O. (2011b). Research on data use: A framework and analysis. *Measurement Interdisciplinary Research and Perspectives, 9*(4), 173–206. https://doi.org/10.1080/15366367.2011b.626729

Coburn, C. E., & Turner, E. O. (2012). The practice of data use: An introduction. *American Journal of Education, 118*(2), 99–111. https://doi.org/10.1086/663272

Cohen, D. K., & Mehta, J. D. (2017). Why reform sometimes succeeds: Understanding the conditions that produce reforms that last. *American Educational Research Journal, 54*(4), 644–690. https://doi.org/10.3102/0002831217700078

Cohen, L., Manion, L., & Morrison, K. (2013). *Research methods in education*. Routledge.

Collinson, V., & Cook, T. F. (2006). *Organizational learning: Improving learning, teaching, and leading in school systems*. Sage.

Creemers, B. P. M., & Kyriakides, L. (2010). Explaining stability and changes in school effectiveness by looking at changes in the functioning of school factors. *School Effectiveness and School Improvement, 21*(4), 409–427. https://doi.org/10.1080/09243453.2010.512795

Cresswell, J. W., & Plano Clark, V. (2011). *Designing and conducting mixed method research* (2nd ed.). Sage Publications.

Daly, A. J., & Finnigan, K. S. (2010). A bridge between worlds: Understanding network structure to understand change strategy. *Journal of Educational Change, 11*(2), 111–138.

Datnow, A. (2005). The sustainability of comprehensive school reform models in changing district and state contexts. *Educational Administration Quarterly, 41*(1), 121–153. https://doi.org/10.1177/0013161X04269578

Datnow, A., & Hubbard, L. (2016). Teacher capacity for and beliefs about data-driven decision making: A literature review of international research. *Journal of Educational Change*. https://doi.org/10.1007/s10833-015-9264-2

Datnow, A., & Park, V. (2015). Data use for equity. *Educational Leadership, 72*(5), 48–54.

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. https://doi.org/10.1108/09578231311311500

Deaney, R., & Hennessy, S. (2007a). Sustainability, evolution and dissemination of information and communication technology-supported classroom practice. *Research Papers in Education*. https://doi.org/10.1080/02671520601152102

Deaney, R., & Hennessy, S. (2007b). Sustainability, evolution and dissemination of information and communication technology-supported classroom practice. *Research Papers in Education, 22*(1), 65–94. https://doi.org/10.1080/02671520601152102

Dekker, T., & Feijs, E. (2005). Scaling up strategies for change: Change in formative assessment practices. *Assessment in Education: Principles, Policy and Practice, 12*(3), 237–254. https://doi.org/10.1080/09695940500337215

Denzin, N. K., & Lincoln, Y. S. (Eds.). (2013). *Collecting and interpreting qualitative materials* (4th ed.). Sage.

Dosi, G., Nelson, R. R., Winter, S. G., et al. (2000). *The nature and dynamics of organizational capabilities*. Oxford University Press.

Drits-Esser, D., Gess-Newsome, J., & Stark, L. A. (2017). Examining the sustainability of teacher learning following a year-long science professional development programme for inservice primary school teachers. *Professional Development in Education, 43*(3), 375–396. https://doi.org/10.1080/19415257.2016.1179664

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. https://doi.org/10.1016/j.stueduc.2015.11.002

Edwards Groves, C., & Rönnerman, K. (2013). Generating leading practices through professional learning. *Professional Development in Education, 39*(1), 122–140. https://doi.org/10.1080/19415257.2012.724439

Elder, K. I., & Prochnow, J. E. (2016). PB4L school-wide: What will support the sustainability of the initiative? *New Zealand Journal of Educational Studies, 51*(1), 83–97. https://doi.org/10.1007/s40841-016-0036-1

Elias, M. (2010). Sustainability of Social-Emotional Learning and related programs: Lessons from a field study. *The International Journal of Emotional Education, 2*(1), 17–33.

Fagen, M. C., & Flay, B. R. (2009). Sustaining a school-based prevention program: Results from the Aban Aya Sustainability Project. *Health Education and Behavior, 36*(1), 9–23. https://doi.org/10.1177/1090198106291376

Feldman, M. S., & Pentland, B. T. (2003b). Reconceptualizing organizational routines as a source of flexibility and change. In *Administrative science quarterly* (Vol. 48, Issue 1, pp. 94–118). Johnson School at Cornell University. https://doi.org/10.2307/3556620

Feldman, M. S., & Pentland, B. T. (2003a). *Reconceptualizing and Change. Adhesives Sealants Industry, 48*(3), 94–118.

Ferguson, N., Currie, L.-A., Paul, M., & Topping, K. (2011). The longitudinal impact of a comprehensive literacy intervention. *Educational Research, 53*(3), 237–256. https://doi.org/10.1080/00131881.2011.598657

Floyd, J., & Fowler, J. (2014). *Survey research methods* (5th ed.). Sage.

Fullan, M. (2007). The new meaning of educational change. In *School effectiveness and school improvement* (5th ed., Vol. 2, Issue 4). Teachers College Press. https://doi.org/10.1080/0924345910020406

Fullan, M. (1992). *Successful school improvement: The implementation perspective and beyond.* McGraw-Hill Education.

Furman Shaharabani, Y., & Tal, T. (2017). Teachers' practice a decade after an extensive professional development program in science education. *Research in Science Education, 47*(5), 1031–1053. https://doi.org/10.1007/s11165-016-9539-5

Gaikhorst, L., Mioch, R., & Weijers, D. (2018). *Een overzichtsstudie. February*.

Gaikhorst, L., Beishuizen, J. J. J., Zijlstra, B. J. H., & Volman, M. L. L. (2017). The sustainability of a teacher professional development programme for beginning urban teachers. *Cambridge Journal of Education, 47*(1), 135–154. https://doi.org/10.1080/0305764X.2015.1125449

Giles, C., & Hargreaves, A. (2006). The sustainability of innovative schools as learning organizations and professional learning communities during standardized reform. *Educational Administration Quarterly, 42*(1), 124–156. https://doi.org/10.1177/0013161X05278189

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis*. Prentice Hall.

Hargreaves, A., & Fink, D. (2000). The three dimensions of reform. *Educational Leadership, 57*(7), 30–34.

Hargreaves, A., & Fink, D. (2012). *Sustainable leadership* (Vol. 6). Wiley.

Hargreaves, A., & Goodson, I. (2006). Educational change over time? The sustainability and nonsustainability of three decades of secondary school change and continuity. *Educational Administration Quarterly, 42*(1), 3–41.

Howard, S. K., Schrum, L., Voogt, J., & Sligte, H. (2021). Designing research to inform sustainability and scalability of digital technology innovations. *Educational Technology Research and Development, 69*(4), 2309–2329. https://doi.org/10.1007/s11423-020-09913-y

Howard-Grenville, J., Rerup, C., Langley, A., & Tsoukas, H. (2016). Organizational routines: How they are created, maintained, and changed. *Perspectives on Process Organization Studies, 6*(1), 1–18.

Hubers, M. D. (2016). *Capacity building by data team members tot sustain schools' data use*.

Hubers, M. D., Moolenaar, N. M., Schildkamp, K., Daly, A. J., Handelzalts, A., & Pieters, J. M. (2018). Share and succeed: the development of knowledge sharing and brokerage in data teams' network structures. *Research Papers in Education, 33*(2), 25. https://doi.org/10.1080/02671522.2017.1286682

Hubers, M. D., Poortman, C. L., Schildkamp, K., & Pieters, J. M. (2019). Spreading the word: Boundary crossers building collective capacity for data use. *Teachers College Record, 121*(1), 1–45.

Hubers, M. D., Schildkamp, K., Poortman, C. L., & Pieters, J. M. (2017). The quest for sustained data use: Developing organizational routines. *Teaching and Teacher Education*. https://doi.org/10.1016/j.tate.2017.07.007

Jacobs, D., & Snijders, H. (2008). *Innovatieroutine Hoe managers herhaalde innovatie kunnen stimuleren*. Van Gorcum.

Jimerson, J. B., Garry, V., Poortman, C. L., & Schildkamp, K. (2020). Implementation of a collaborative data use model in a United States context. *Studies in Educational Evaluation*. https://doi.org/10.1016/j.stueduc.2020.100866

Jimerson, J. B., & Wayman, J. C. (2015). Professional learning for using data: Examining teacher needs and supports. *Teachers College Record, 117*(4), 1–36.

Kafyulilo, A., Fisser, P., & Voogt, J. (2016). Factors affecting teachers' continuation of technology use in teaching. *Education and Information Technologies, 21*(6), 1535–1554. https://doi.org/10.1007/s10639-015-9398-0

King, F. (2016). Teacher professional development to support teacher professional learning: Systemic Factors from Irish case studies. *Teacher Development, 20*(4), 574–594. https://doi.org/10.1080/13664530.2016.1161661

Kippers, W. B., Poortman, C. L., Schildkamp, K., & Visscher, A. J. (2018). Data literacy: What do educators learn and struggle with during a data use intervention? *Studies in Educational Evaluation, 56*(September 2017), 21–31. https://doi.org/10.1016/j.stueduc.2017.11.001

Kremser, W., & Schreyögg, G. (2016). The dynamics of interrelated routines: introducing the cluster level. *Organization Science*. https://doi.org/10.1287/orsc.2015.1042

Lai, M. K., & McNaughton, S. (2016). The impact of data use professional development on student achievement. *Teaching and Teacher Education*. https://doi.org/10.1016/j.tate.2016.07.005

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* (pp. 9–21). Springer. https://doi.org/10.1007/978-94-007-4816-3

Larsen, T., & Samdal, O. (2008). Facilitating the implementation and sustainability of second step. *Scandinavian Journal of Educational Research, 52*(2), 187–204. https://doi.org/10.1080/00313830801915820

Lee, M., & Louis, K. S. (2019). Mapping a strong school culture and linking it to sustainable school improvement. *Teaching and Teacher Education, 81*, 84–96. https://doi.org/10.1016/j.tate.2019.02.001

Mandinach, E. B., Friedman, J. M., & Gummer, E. S. (2015). How can schools of education help to build educators' capacity to use data? A systemic view of the issue. *Teachers College Record, 117*(4), 1–50.

Mandinach, E. B., & Gummer, E. S. (2013a). A systemic view of implementing data literacy in educator preparation. *Educational Researcher, 42*(1), 30–37. https://doi.org/10.3102/0013189X12459803

Marsh, J. A. (2012b). *Teachers College Record, 14* (11), 1–65.

Marsh, J. (2012a). Interventions promoting educators' use of data: Research insights and gaps. *Teachers College Record, 114*(11), 1–48.

Marsh, J. A. (2016). The political dynamics of district reform: The form and fate of the los angeles public school choice initiative. *Teachers College Record, 118*(9), 1–40. https://doi.org/10.1017/CBO9781107415324.004

März, V., Geijsel, F., & März, V. (n.d.). *Van acties naar interacties*.

Mcnaughton, S. (2021). *The Conundrum Research-Practice Partnerships Face with System Variability*. https://doi.org/10.1016/j.stueduc.2021.101048

Mijs, D. (2007). *Effectieve Schoolverbetering*.

Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. (2007). *Grondwet voor het Koninkrijk der Nederlanden (1814-)*. http://deeplinking.kluwer.nl/docid/inod3c86b101d4fdbda1269a38a81dccc70a

Mitra, D. L. (2009). The role of intermediary organizations in sustaining student voice initiatives. *Teachers College Record, 111*(7), 1834–1869.

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health and Mental Health Services Research, 42*(5), 533–544. https://doi.org/10.1007/s10488-013-0528-y

Patton, M. (2002). *Qualitative research and evaluation methods*.

Payneeandy, S. (2012). *School-based teacher training and the development of literacy in low-achieving schools school-based training*. *18*(10).

Pluye, P., Potvin, L., Denis, J. L., Pelletier, J., & Mannoni, C. (2005). Program sustainability begins with the first events. *Evaluation and Program Planning, 28*(2), 123–137. https://doi.org/10.1016/j.evalprogplan.2004.10.003

Poortman, C. L., & Schildkamp, K. (2016). Solving student achievement problems with a data use intervention for teachers. *Teaching and Teacher Education, 60*, 425–433. https://doi.org/10.1016/j.tate.2016.06.010

Prenger, R., Tappel, A.P.M., Poortman, C. L., & Schildkamp, K. (2022, in press.). *How can educational innovations become sustainable? A review of empirical literature*.

Roffe, I. (2010). Sustainability of curriculum development for enterprise education: Observations on cases from Wales. *Education + Training, 52*(2), 140–164. https://doi.org/10.1108/00400911011027734

Saito, E., Khong, T. D. H., & Tsukui, A. (2012). Why is school reform sustained even after a project? A case study of Bac Giang Province. *Vietnam. Journal of Educational Change, 13*(2), 259–287. https://doi.org/10.1007/s10833-011-9173-y

Sanders, M. G. (2012a). Achieving scale at the district level: A longitudinal multiple case study of a partnership reform. *Educational Administration Quarterly, 48*(1), 154–186. https://doi.org/10.1177/0013161X11417432

Sanders, M. G. (2012b). Sustaining programs of school, family, and community partnerships. *Educational Policy, 26*(6), 845–869. https://doi.org/10.1177/0895904811417591

Schildkamp, K., Handelzalts, A., Poortman, C. L., Leusink, H., Meerdink, M., Smit, M., Ebbeler, J., & Hubers, M. D. (2014). *De datateam methode: Een concrete aanpak voor onderwijsverbetering* (1st ed.). Maklu.

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. https://doi.org/10.1016/j.tate.2009.06.007

Schildkamp, K., & Lai, M. K. (2013). Data-based decision making in education: Challenges and opportunities. In K. Schildkamp, M. K. Lai, & L. Earl (Eds.), *Introduction* (pp. 1–7). Springer.

Schildkamp, K., & Poortman, C. L. (2015). Factors influencing the functioning of data teams. *Teachers College Record, 117*(4), 1–42.

Schildkamp, K., Poortman, C. L., & Handelzalts, A. (2016). Data teams for school improvement. *School Effectiveness and School Improvement, 27*(2), 228–254. https://doi.org/10.1080/09243453.2015.1056192

Schildkamp, K., Smit, M., & Blossing, U. (2019). Professional development in the use of data: From data to knowledge in data teams. *Scandinavian Journal of Educational Research, 63*(3), 393–411. https://doi.org/10.1080/00313831.2017.1376350

Sherer, J. Z., & Spillane, J. (2011). Constancy and change in work practice in schools: The role of organizational routines. *Teachers College Record, 113*(3), 611–657.

Stringfield, S., Reynolds, D., & Schaffer, E. C. (2008). Improving secondary students' academic achievement through a focus on reform reliability: 4- and 9-year findings from the High Reliability Schools project. *School Effectiveness and School Improvement, 19*(4), 409–428. https://doi.org/10.1080/09243450802535190

Tam, F. W. (2009). Sufficient conditions for sustainable instructional changes in the classroom: The case of Hong Kong. *Journal of Educational Change, 10*(4), 315–336. https://doi.org/10.1007/s10833-008-9091-9

van den Boom-Muilenburg, S. N. (2021). *The role of school leadership in schools that work sustainably on school improvement with professional learning communities*. University of Twente.

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 in primary schools. *American Educational Research Journal, 53*(2), 360–394. https://doi.org/10.3102/0002831216637346

van Geel, M., Keuning, T., Visscher, A., & Fox, J. P. (2017). Changes in educators' data literacy during a data-based decision making intervention. *Teaching and Teacher Education, 64*, 187–198. https://doi.org/10.1016/j.tate.2017.02.015

Van Voorhis, F., & Sheldon, S. (2004). Principals' roles in the development of US programs of school, family, and community partnerships. *International Journal of Educational Research, 41*(1 SPEC. ISS.), 55–70. https://doi.org/10.1016/j.ijer.2005.04.005

van Veen, K., Zwart, R., Meirink, J., & Verloop, N. (2010). Professionele ontwikkeling van leraren: een reviewstudie naar effectieve kenmerken van professionaliseringsinterventies van leraren. *ICLON/Expertisecentrum Leren van Docenten* (December), 2/150. http://www.nro.nl/wp-content/uploads/2014/05/PROO+Professionele+ontwikkeling+van+leraren+Klaas+van+Veen+ea.pdf

Venkat, H., & Adler, J. (2008). Expanding the foci of activity theory: Accessing the broader contexts and experiences of mathematics education reform. *Educational Review, 60*(2), 127–140.

Visscher, A. J., & Ehren, M. C. M. (2011). *De eenvoud en complexiteit van Opbrengstgericht Werken*. Universiteit Twente.

Waslander, S. (2007). Leren over innoveren. Overzichtsstudie van wetenschappelijk onderzoek naar duurzaam. *VO-Project Innovatie*.

Wiltsey-Stirrnam, S., et al. (2012). The sustainability of new programs and innovations: A review of the empirical literature and recommendations for future research. *Implementation Science, 7*(17), 1–19.

Wolthuis, F., van Veen, K., de Vries, S., & Hubers, M. D. (2020). Between lethal and local adaptation: Lesson study as an organizational routine. *International Journal of Educational Research, 100*(July 2019), 101534. https://doi.org/10.1016/j.ijer.2020.101534

Wood, P. (2017). Overcoming the problem of embedding change in educational organizations: a perspective from normalization process theory. *Management in Education, 31*(1), 33–38. https://doi.org/10.1177/0892020616685286