Chapter 11
Recurrence Quantification Analysis as a Methodological Innovation for School Improvement Research

Arnoud Oude Groote Beverborg, Maarten Wijnants, Peter J. C. Sleegers, and Tobias Feldhoff

11.1 Introduction

In educational research and practice, teacher learning in schools is recognized as an important resource in support of school improvement and educational change. In their efforts to understand the mechanisms underlying school improvement, researchers have started to examine the role of teacher learning as a key component to building school-wide capacity to change. In practice, professional learning communities are being increasingly developed to stimulate the sharing of knowledge, information and expertise among teachers, with the goal to improve instruction and student learning. More specifically, by engaging in professional learning activities, teachers can make knowledge and information explicit, discover the proper scripts for future actions aimed at adaptation to changes such as ongoing reorganizations of work processes and accountability reforms, and to formulate and monitor goals for further development of for instance instructional methods and technological innovations (Korthagen, 2010; Oude Groote Beverborg, Sleegers, Endedijk, & van Veen, 2015a).

To understand more about how engagement in professional learning activities enables teachers to learn, scholars have called for more situated and longitudinal research (Feldhoff, Radisch, & Bischof, 2016; Feldhoff, Radisch, & Klieme, 2014; Korthagen, 2010). The few longitudinal studies conducted so far used analytic

A. Oude Groote Beverborg (✉) · M. Wijnants
Radboud University Nijmegen, Nijmegen, The Netherlands
e-mail: a.oudegrootebeverborg@fm.ru.nl

P. J. C. Sleegers
BMC, Amersfoort, The Netherlands

T. Feldhoff
Johannes Gutenberg University, Mainz, Germany

219

techniques (Structural Equation Modelling; SEM) that derive their power from large samples of participants and included a limited number of measurement occasions with relatively long intervals (e.g. yearly intervals) to assess the (reciprocal) relationships between variables under study. The findings suggest, among other things, that reflection is positively related to self-efficacy and changes in instructional practices (Oude Groote Beverborg, et al., 2015a; Sleegers, Thoonen, Oort, & Peetsma, 2014). Higher levels of engagement in professional learning activities, thus, seem beneficial to improve education. In addition, these studies pointed towards the importance of conditions at the school-level, such as transformational leadership and working in teams, to foster teacher learning. This suggests that a purposeful and empowering environment can help to structure uncertainty and ambiguity, and to enable teachers to come to a common understanding about changing their practice, and learn from one another (see also Coburn, 2004; Oude Groote Beverborg, 2015; Staples & Webster, 2008). As such, these longitudinal studies have their merit in validating and extending previous findings from cross-sectional studies on the structural relations between organizational conditions and improving education over time (see also Hallinger & Heck, 2011; Heck & Hallinger, 2009; Heck & Hallinger, 2010).

However, findings on structures at the school-level do not inform about how teachers use these organizational conditions in everyday regulation practices and how such use may fluctuate over time (Maag Merki, Grob, Rechsteiner, Rickenbacher, & Wullschleger, 2021, see chapter 12; see also Hamaker, 2012; Molenaar & Campbell, 2009). It remains for instance unclear how higher levels of engagement in professional learning activities translate to individual teachers' routines of for instance reflection or knowledge sharing on a daily basis (see also Little & Horn, 2007). Are these higher levels based on for instance reflecting very regularly (every day a little) or in bursts (whenever there is a necessity or opportunity)? By extension, it remains unclear whether the regularity with which moments of teacher learning are organized also contributes to sustaining school improvement (think with regard to regularity for instance of the rhythm of reflection cycle phases for self-improvement, the periodicity of meetings of learning community members to develop instruction and curriculum, and even the intervals of appraisal interviews and classroom observations that can be used for quality development monitoring and accountability purposes) (e.g. Desimone, 2009; Korthagen, 2001; van der Lans, 2018; van der Lans, van de Grift, & van Veen, 2018).

In contrast to large survey studies, case studies have generated situated descriptions of what occurs during efforts to improve schools in specific contexts (see for instance Coburn, 2001, 2005, 2006). However, case studies do not have the aim to generalize their findings, and the validity and utility of those findings is limited. As such, the available research provides no systematic evidence of how (for what and when) teacher learning takes shape in its social context. Consequently, understanding more about the dynamics of everyday teacher learning and its link with school improvement and educational change requires studies that are situated, longitudinal,

and aimed at finding systematic relations, and in addition, a corresponding situated and dynamic perspective (Barab et al., 1999; Clarke & Hollingsworth, 2002; Greeno, 1998; Heft, 2001; Horn, 2005; Lave & Wenger, 1991; Reed, 1996).

From a situated and dynamic perspective, school improvement is seen as an ongoing, embedded, complex, and dynamic process of adapting to continuously changing challenges that arise out of schools' unique circumstances. School improvement emerges from the many interactions between actors within and outside schools, making the school improvement journey highly context-sensitive, and the occurrence of meaningful developments (or milestones) unpredictable (van Geert & Steenbeek, 2014; see also Ng, 2021, chapter 7). Similarly, teacher learning is seen as a cyclical process in which available environmental information, professional learning activities, and productive practices are interconnected and co-develop (Barab et al., 1999; Clarke & Hollingsworth, 2002), that is, teachers attend to, interpret, adapt, and transform information from their environment and make use of their (social) environment to learn what is needed (Barab & Roth, 2006; Gibson, 1979/1986; Greeno, 1998; Little, 1990; Maitlis, 2005).

Investigating ongoing micro-level change processes, such as the routine with which individual teachers make environmental information and changes in meaning, knowledge, or accommodation of teaching practices, explicit, requires analytic techniques that assess intra-individual variability over time, such as State Space Grid analysis (Granic & Dishion, 2003; Lewis, Lamey, & Douglas, 1999; Mainhard, Pennings, Wubbels, & Brekelmans, 2012) or Recurrent Quantification Analysis (RQA). In contrast to commonly used statistical modelling techniques, such as SEM, these techniques are based on dense time-series, whose temporal structures are kept intact. They provide measures about the stability or flexibility of a developmental process. RQA has been applied to analyse coordination in conversations, reading fluency, emergence of insights and behavioural changes (Dale & Spivey, 2005; Lichtwarck-Aschoff, Hasselman, Cox, Pepler, & Granic, 2012; O'Brien, Wallot, Haussmann, & Kloos, 2014; Richardson, Dale, & Kirkham, 2007; Stephen, Dixon, & Isenhower, 2009; Wijnants, Hasselman, Cox, Bosman, & Van Orden, 2012; see also Wijnants, Bosman, Hasselman, Cox, & Van Orden, 2009).

This study aims to examine the overall level and the routine of learning through reflection in the workplace. More specifically, this study focusses on the relation between the temporal pattern of becoming aware of information in the (social) environment and experiencing new insights by making both explicit through reflection. It does so by collecting dense intra-individual (teacher) longitudinal measurements (logs), and by illustrating how RQA can be applied to these time-series. We will explore the application of RQA as a promising analytic technique for understanding the co-evolution of teacher learning and school-wide capacity for sustained improvement.

11.2   Theoretical and Methodological Framework

In this section, we will first describe teachers as active interpreters of their specific circumstances and as reflective practitioners (e.g. Clarke & Hollingsworth, 2002). Next, we will discuss and describe logs as measurement instruments that can capture this situated process over time. Thereafter, we will extensively discuss RQA and we will present examples of studies to provide some research context as to how it can be applied. We will end this section by showing how this conceptualization, measurement instrument, and analysis strategy come together in the present study.

11.2.1   *Information and Reflection in a Situated and Ongoing Learning Process*

Within the situated perspective, teacher learning is considered an acculturation process (Greeno, 1998; Lave & Wenger, 1991). Teachers are considered active, intentional perceivers, constructing a meaningful practice by integrating new experiences with old experiences (Coburn, 2004; Sleegers & Spillane, 2009; Spillane & Miele, 2007). These experiences are provided by the community while the person is engaged in it (Lave & Wenger, 1991; Little, 2003; Wenger, 1998). Central to this perspective is that knowledge is distributed over a situation (Greeno, 1998; Hutchins, 1995; Putnam & Borko, 2000), that a person makes sense of it through action (Little, 2003; Spillane, Reiser, & Reimer, 2002; Weick, 2011), and that sensemaking is embedded in a person's history (Coburn, 2001; Coburn, 2004; Sleegers, Wassink, van Veen, & Imants, 2009), as well as in a social and cultural context (Sleegers & Spillane, 2009). While acting, a person selects the information that affords continued action and that fits the understanding of the purpose in the situation (Coburn, 2001; Sleegers et al., 2009; Spillane et al., 2002). Learning can thereby also be characterized as a process of continuously attuning (Barab et al., 1999; Clarke & Hollingsworth, 2002; Granic & Dishion, 2003; Guastello, 2002). As such, teachers can regulate what information in the (social) environment they attend to, so that, over a longer period of time, experiences of interactions with the (social) environment consolidate into new, or differentiations of, meanings, knowledge, and skills (Korthagen, 2010; Kunnen & Bosma, 2000; Lichtwarck-Aschoff, Kunnen, & van Geert, 2009; Steenbeek & van Geert, 2007; van Geert & Steenbeek, 2005). In addition, of course, teachers can develop and adapt by regulating their activities through reflection (Argyris & Schön, 1974; Korthagen & Vasalos, 2005; Schön, 1983).

Teacher engagement in reflection, then, can be seen as an introspective activity that refers to a person recreating an experience of acting in a given situation. In making this experience explicit later, a person supplements the memory of the experience with new ideas that can either be self-generated or based on information gained from others (Oude Groote Beverborg, Sleegers, & van Veen, 2015b). This creates an altered and thus new experience, which can then serve as the basis for

future action. In this way, reflection directs what information in the environment is to be attended to, thought about, and reacted to, and for what purpose (Clarke & Hollingsworth, 2002; see also Weick, 2006). Making information explicit in this way helps to put the knowledge that is distributed within teachers' environments to focussed use and regulates development and adaptation by setting priorities for attention and actions. As such, making previously encountered information explicit shapes future experiences, what can be consequently reflection upon, and what will be made explicit thereafter. This interplay between environmental information and reflection stresses that the directions teachers' and their school's developments can take are based in a teacher's specific circumstances.

Moreover, through repeated investigation of one's own actions and encountered information, a teacher might, after a while, suddenly discover a new way of acting or looking at the world that is more functional in a given situation than the old one was (Clarke & Hollingsworth, 2002). Such learning experiences of change in meaning, knowledge, or skills, which were generated by one person, can also be reflected upon, made explicit, and shared as possibly of value for other individuals and the team (Nonaka, 1994; van Woerkom, 2004). That also helps to find solutions to ongoing changes and challenges at work, and to formulate and monitor goals for further development (of for instance shared meaning) and improvement (of for instance a school's capacity for change) (Oude Groote Beverborg, et al., 2015a).

However, due to the circumstantial and temporal dependency of available information, meaning, knowledge, and skills, intensities of engagement in reflection on one's working environment can fluctuate over time within persons and can differ between persons before new insights emerge (Endedijk, Brekelmans, Verloop, Sleegers, & Vermunt, 2014; Stephen & Dixon, 2009; see also Orton & Weick, 1990). The corresponding trajectories of individual teachers' engagements in making information explicit may therefore look quite irregular and not alike. Additionally, learning experiences can also emerge with different intervals. Repeated engagement in reflection on one's working environment therefore changes, continuously slightly (sensitivity to specific information) and occasionally more profoundly (experience of having learned something), the way the world is perceived, understood, and enacted (see also Coburn, 2004; Voestermans & Verheggen, 2007, 2013).

Nevertheless, it remains unclear with how much routine teachers engage in reflection in their everyday practices. Insights into the intra-individual variability in intensity of everyday reflection may provide valuable knowledge to schools as well as to the inspectorates of education about the ways, in which they can organize and support teacher learning in the workplace. In order to tap into these dynamics of reflection and their consequences, measurement instruments therefore need to be designed that allow for specific person-environment interactions and that can be administered densely (see also Bolger & Laurenceau, 2013). Moreover, the chosen analysis needs to provide measures that can represent temporal variability. In the next two sections, we will address the use of logs as a measurement instrument that can be administered densely and the use of RQA as an analytic technique that yields dynamics measures.

11.2.2   Logs

In order to tap into the dynamics of individual teachers' reflection processes, it is necessary to look at them while and where they are happening – rather than by means of for instance interviews that are prone to hindsight bias or with standardized questionnaires that are insensitive to specific circumstances – to focus on the continuous interaction between the acting professional and the environment through time, and then reconstruct the learning process as a series of interactions over time (see for an example Endedijk, Hoekman, & Sleegers, 2014; Lunenberg, Korthagen, & Zwart, 2011; Lunenberg, Zwart, & Korthagen, 2010; Zwart, Wubbels, Bergen, & Bolhuis, 2007; Zwart, Wubbels, Bolhuis, & Bergen, 2008). This would give an account of professional development including prospective learning, and not only an account of retrospective learning.

In this study, we will therefore measure teachers' reflection processes with logs (for other uses of logs in dynamic analyses, see: Guastello, Johnson, & Rieke, 1999; Lichtwarck-Aschoff et al., 2009; Maitlis, 2005; for other uses of logs in school improvement research, see Maag Merki et al., 2021, chapter 12; Spillane & Zuberi, 2021, chapter 9). Not everything that happens can be reported in a log. What is reported, is what is most salient in a teacher's experience. Using open questions, this can be charted in a personalized and situated manner.

The use of logs presupposes that teachers have a sensitivity to information in their environment, that they monitor their development, and that they have an affinity for making information and knowledge explicit by using logs. Every time teachers fill in a log entry, they use an opportunity to make information, experiences, or knowledge explicit (as, in a sense, surveys with targeted items and interviews with targeted questions do as well). Participating in this study might therefore make teachers more aware of what is going on in their environment, of their purpose, and in what areas they develop (Geursen, de Heer, Korthagen, Lunenberg, & Zwart, 2010). By administering logs densely, the logs themselves can also become a familiar part of the working environment that teachers can choose to engage with. Nevertheless, teachers flow with the issues of the day, and may find it hard to disengage from the immediacy of their work to make time to reflect by using logs. Logs thereby not only measure the learning process. They do so by setting a model of the reflection process in terms of content and pace that may fit better or worse to different teachers within a certain period of time. Moreover, the interval with which logs are administered ought to be in accord with the expected rate of change of the frequency with which teachers are likely to reflect upon their environment and learning experiences.

For the assessment of reflection routines, it is important that logs can generate a dense time-series. From these time-series, the dynamics of engagement in reflection can be reconstructed with an RQA.

11.2.3 Recurrence Quantification Analysis

RQA is a nonlinear technique to quantify recurring patterns and parameters pertaining to the stability of the underlying dynamics from a time-series (with an intact temporal structure). An important advantage of RQA, unlike other time-series analysis methods, is that this technique does not impose constraints on data-set size (N). RQA does not make assumptions regarding statistical distributions or stationarity of data either. Nevertheless, for RQA to provide interpretable results, it has been suggested that the minimum requirements for the time-series are that they are long enough to contain at least two repetitions of the whole repeating dynamic pattern and that at least three measurement occasions fall within each repetition of the repeating dynamic pattern (Brick, Gray, & Staples, 2018). Needless to say that more robust and precise estimation will be permitted by measuring longer and denser, which may thus be required for noisier data. The technique reveals subtle time-evolutionary behaviour of complex systems by quantifying system characteristics that would otherwise have remained hidden (i.e., when only taking frequencies into account). To get an idea of what is meant by dynamics, consider Fig. 11.1. It shows five examples of hypothetical, idealized change trajectories (i.e. stability, growth, randomness, and two times regular fluctuation) of engagement in reflection of different persons. Trajectories (**a, b**, **c,** and **d**) all have different temporal patterns (rhythms). Their overall level of reflection does not distinguish them: Each trajectory has a mean of 1. In comparison, trajectories (**d** and **e**) differ in their means, but have the same rhythm. The differences between the change trajectories become apparent, because they have (relatively) many time-points.

A distinction can be made between the application of RQA to categorical (nominal) data[1] and to continuous (scale) data. Categorical RQA is a simplified form of continuous RQA[2]. This chapter will focus on categorical RQA. Moreover, RQA can be applied to single time-series (auto-RQA) or to two different time-series (cross-RQA). Fundamentally, auto-RQA is applied to answer questions concerning

---

[1] RQA allows a direct access to dynamic systems (characterized by a large number of participating, often interacting variables) by reconstructing, from a single measured variable in the interactive system, a behaviour space (or phase-space) that represents the dynamics of the entire system. This reconstruction is achieved by the method of delay-embedding that is based on Takens' theorem (Broer & Takens, 2009; Takens, 1981). The phase space reconstructed from the time series of this single variable informs about the behaviour of the entire system because the influence of any interdependent, dynamical variable is contained in the measured signal. The reconstruction itself involves creating time-delayed copies of the time-series of a variable that become the surrogate dimensions of a multi-dimensional phase-space. Consequently, the original variable becomes a dimension of the system in question and each time-delayed copy becomes another dimension of the system. Because of that, it is not needed to know all elements of the system, or measure them, to reconstruct the behaviour of a dynamic system, provided that a (sufficiently dense) time-series of one element of the system is available. For tutorials on continuous RQA, see: Marwan et al. (2007) and Riley and Van Orden (2005). For applications of continuous RQA in the social sciences, see: Richardson, Schmidt, and Kay (2007) and Shockley, Santana, and Fowler (2003).

[2] Delay-embedding is not applied – the system is considered to have 1 dimension.

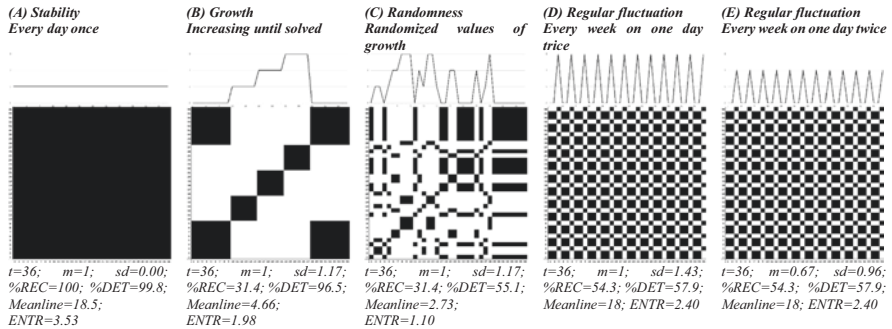| (A) Stability Every day once	(B) Growth Increasing until solved	(C) Randomness Randomized values of growth	(D) Regular fluctuation Every week on one day trice	(E) Regular fluctuation Every week on one day twice
t=36; m=1; sd=0.00; %REC=100; %DET=99.8; Meanline=18.5; ENTR=3.53	t=36; m=1; sd=1.17; %REC=31.4; %DET=96.5; Meanline=4.66; ENTR=1.98	t=36; m=1; sd=1.17; %REC=31.4; %DET=55.1; Meanline=2.73; ENTR=1.10	t=36; m=1; sd=1.43; %REC=54.3; %DET=57.9; Meanline=18; ENTR=2.40	t=36; m=0.67; sd=0.96; %REC=54.3; %DET=57.9; Meanline=18; ENTR=2.40

Fig. 11.1 Five examples of change trajectories, shown as time-series graphs and recurrence plots, of engagement in reflection with different dynamics

Note: Change trajectories (**a**, **b**, **c**, **d** and **e**) represent hypothetical, idealized change trajectories (i.e. stability, growth, randomness, and two times regular fluctuation, respectively) of engagement in reflection of different persons. Trajectories (**a**, **b**, **c** and **d**) all have a mean of 1 but differ in the values of their dynamics (rhythm) measures. In comparison, trajectories (**d** and **e**) differ in their means, but have the same values of their dynamics measures. Each trajectory is represented by two graphs: one time-series and one recurrence plot (top and bottom graphs, respectively). The time-series have 36 time points (i.e. days) (x-axis of each graph) and engagement in reflection can have one of the following values at each time point: 0, 1, 2, or 3 (i.e, the number of reflection moments, or the amount of reflection intensity, per day) (y-axis of each graph). In the recurrence plots, both the x-axis and the y-axis represent the 36 time points, and thus the plots have 36*36 = 1296 cells. These cells can either be filled or empty (filling is in this case marked by a black square). Filled cells are called recurrence points. Recurrence points represent that the process had a value at a certain time point and that that value also occurred at another time point (i.e. the recurrence of one of the reflection intensity values). In these examples, the time-series are plotted against themselves in the recurrence plots (i.e. auto-recurrence), and thus the plots are symmetrical around the Line of Incidence (the center diagonal line, i.e. the time-series as it was measured). Auto-recurrence plots are generated for each single time-series separately. The Line of Incidence is excluded in the calculation of the dynamics measures. t = length of the time-series; m = mean of the values in the time-series; sd = standard deviation around the mean; %REC = Recurrence Rate (i.e. the percentage of recurrence points in the recurrence plot); %DET = Determinism (i.e. the percentage of recurrence points that form diagonal lines out of the total of recurrence points); Meanline = the mean length of all diagonal lines of recurrence points; ENTR = Shannon Entropy (i.e. a measure of complexity; it is calculated as the sum of the probability of observing a diagonal Line Length times the log base 2 of that probability). See also the Recurrence Quantification Analysis-section and Fig. 11.3

within-actor variability, whereas cross-RQA is applied to answer questions concerning variability in coordination between actors over time.

RQA combines the visualization of temporal dynamics in recurrence plots with the objective quantification of (non-linear) system properties. In auto-RQA, one time-series is placed on both the x-axis and the y-axis to generate the recurrence plot. In cross-RQA, one time-series is placed on the x-axis and another time-series is placed on the y-axis to generate the recurrence plot. In essence, a recurrence plot is a graphical representation of a binomial matrix that shows after what delays

values in time-series recur (recurrence points[3]). The recurrence plot is then quantified and used to calculate complexity measures.

Consider Fig. 11.1 again. In the figure, engagement in reflection has one of the following values at each time point: 0, 1, 2, or 3 (i.e. the number of reflection moments, or the amount of reflection intensity, per day). The temporal order of these values is given in the time-series graphs. The recurrence plots on the other hand are composed of auto-recurrence points; that is, they show that any of these values occurred at a certain moment and that that also happened sometime else within the same time-series (earlier, at the same time, or later). In these examples, the time-series are plotted against themselves in the recurrence plots (i.e. auto-recurrence), and thus the plots are symmetrical around the Line of Incidence (the centre diagonal line, i.e. the actual time-series – in cross-RQA, this line is sometimes called the Line of Synchrony). Auto-recurrence plots are generated separately for each single time-series. The time-series graph of the stable process in (**a**) shows that at each time point the process had a value of 1. Therefore, the corresponding recurrence plot is fully filled. In comparison, the growth (and decline) process in (**b**) shows a steady increase from 0 to 3 followed by a sharp decrease to 0 again. Consequently, the recurrence plot shows neatly clustered recurrence points. The random process in (**c**) has the same time-series values as the time-series in (**b**), but in (**c**), the temporal structure of these values was changed by placing them in a random order. Consequently, the recurrence plot of the process in (**c**) is less characterized by diagonal lines (consecutive recurrences form diagonal lines). Therefore, the process in (**c**) has the same values as in (**b**) for the mean and the Recurrence Rate, but the other dynamics measures differ. The regularly fluctuating processes in (**d** and **e**) both have only two values (0 and 3, or 0 and 2, respectively), and in both trajectories, these values recur after the same period. Therefore, they have identical recurrence plots and thus identical dynamics measures.

When the same behaviour is repeated periodically or when different behaviours succeed each other periodically, diagonal lines are formed in the recurrence plot. Measures based on the temporal order of these recurrence-sequences in the recurrence plot inform about the dynamics of the system. The Line of Incidence is excluded in the calculation of the dynamics measures. We will introduce the measures Recurrence Rate, Determinism, Meanline, and Entropy (other measures are Maxline, Laminarity, and Trapping Time) (Marwan, Romano, Thiel, & Kurths, 2007; see also Cox, van der Steen, Guevara, de Jonge-Hoekstra, & van Dijk, 2016) and elaborate on three studies as examples of how to apply them.

Recurrence Rate is computed as the ratio of the number of recurrent points (the black regions in the recurrence plot) over the total number of possible recurrence points in the recurrence plot (i.e. the length of the time-series squared). The Recurrence Rate thus indicates how often behaviours in a time-series re-occur (or also occur in the case of cross-RQA). The Recurrence Rate is not based on the

---

[3] Note that for categorical RQA, values need to be clearly demarcated categories to form recurrence points.

temporal order of the values in the time-series, and is thus a raw measure of variability of behaviour (or of coordination in the behaviours of two actors in the case of cross-RQA) over time.

Determinism is defined as the ratio of the number of recurrence points forming a diagonal pattern (i.e. a sequence of recurring behaviours) over the total number of recurrence points in the recurrence plot. Determinism thus informs about behaviours that continue to recur over time relative to isolated recurrences, indicating the persistence of those behaviours.

An example of a study using Recurrence Rate and Determinism was conducted by Dale and Spivey (2005). They applied categorical cross-RQA to assess lexical and syntactic coordination in conversations of dyads of children and caregivers at many measurement occasions (N_{dyads} = 3; $N_{participants}$ = 6; $N_{conversations}$ were 181, 269, and 415). They used the Recurrence Rates of words and of grammar as an indication of coordination between child and caregiver. Types of words are more numerous in conversations than syntactic classes, and types of words therefore gives lower Recurrence Rate values. Additionally, they used the Determinism of words and of grammar, but now based on the set of words that lay within about 50 words from each other in the conversations (i.e. within the band of about 50 words around the Line of Synchrony). This provides an indication of dynamic structures of coordination that are closer together in time and it forms a basis for the interpretation of the Recurrence Rate. Then, they computed both measures again, but now based on the child's time-series at the same measurement occasion and the caregiver's time-series at a measurement occasion one step ahead in development. They compared the 2 × 2 Recurrence measures and the 2 × 2 Determinism measures of each dyad using t-tests to assess the influence of the given conversation. Finally, they assessed the development of the Recurrence Rate and Determinism over time using regression analyses. For all comparisons of RQA measures, results indicated that coordination between child and caregiver was stronger within the same entire conversation than over conversations, and that coordination was stronger with greater temporal proximity within a conversation. Moreover, the results indicated that coordination diminished over development.

Meanline is an index of the average duration of deterministic patterns, and thus indicates how long on average the person (or dyad in the case of cross-RQA) remains in similar behavioural states over time. Meanline provides information about the stability of behaviour.

An example of a study using Meanline was conducted by O'Brien et al. (2014). They applied continuous auto-RQA to assess stability of reading fluency of children in different grades and that of adults ($N_{cohorts}$ = 4; $N_{participants}$ = 71; N_{texts} = 1). All participants read the same text. Additionally, each participant of each cohort was randomly assigned to either a silent reading or a reading out loud condition. The researchers used Meanline as a measure of the length of recurring stretches of word-reading-times (other measures relating to other aspects of reading were also used). ANOVAs were used to compare cohorts and conditions. Moreover, they applied continuous cross-RQA to each possible combination of two time-series of the participants within each cohort and within either condition. This analysis gave

shared-Meanline values. With this measure, an assessment could be made of whether the reading dynamics of each group were more structured by the text (higher shared-Meanline) or more idiosyncratically (lower shared-Meanline), that is, whether more fluent readers are less constrained by the processing of each (subsequent) word and instead follow their own meanderings through the story to monitor their own understanding of the text. Because of concerns that using the pairwise cross-RQA metric may violate the assumption of independence of observations, the shared-Meanline values were submitted to a bootstrap procedure that drew 1000 subsamples per group, after which confidence intervals were constructed for each group. Using 99% confidence intervals, those groups, whose confidence intervals did not overlap, differed significantly from the other groups. The results indicated that adults had more stability in reading in both reading modes as compared to the other cohorts, and that, when reading out loud, the reading dynamics of both sixth graders and adults are structured more idiosyncratically than those of second and fourth graders and also than those of all cohorts during silent reading.

Entropy is computed as the Shannon Entropy of the distribution of the different lengths of the deterministic segments[4] . Entropy indicates the level of complexity of the sequences of behaviours. The Entropy measure, in RQA, thus indicates how much "disorder" there is in the duration of recurrent sequences.

In the form of peak-Entropy, Entropy can for instance be used as a measure of reorganization[5] . Lichtwarck-Aschoff et al. (2012) conducted a study on the course and effect of clinical treatment for externalizing behaviour problems of children (age-range = 7–12 years). A pattern of reorganization over the course of treatment would be an indication of improvement. Both parents and children received treatment once a week for 12 weeks. Bi-weekly 4 or 6-min observations of problem solving discussions between parent and child formed the raw data (N_{dyads} = 41; $N_{participants}$ = 82; $N_{conversations}$ = 6). The data of each participant were initially coded in real-time along nine mutually exclusive affect codes for each participant. The thus acquired time-series were collapsed into one time-series per dyad, resampled to have 72 data points, and recoded along four categories (plus a rest category) that reflected the affective state of the dyad (unordered categorical data). The researchers applied categorical auto-RQA to these dyadic time-series to calculate the Entropy of each conversation of each dyad. 15,000 bootstrap replications of the sample's Entropy values were used to estimate 95% confidence intervals. The

---

[4] Shannon Entropy is calculated as the sum of the probability of observing a diagonal Line Length times the log base 2 of that probability. This measure depends therefore on the number of different lengths of diagonal lines (or bins) in a particular recurrence plot. Fewer bins and more equally distributed frequencies of diagonal Line Lengths over the bins will give lower Entropy values: less information is needed to describe the behaviour of a system.

[5] For instance, learning new knowledge or skills is a reorganization of the (learner's) system in such a way that it becomes (locally) more adapted to its environment. Having learned something new can therefore be characterized by a drop in Entropy, which then stabilizes at this lower level. The reorganization of one's knowledge or skills, on the other hand, is a period, in which old knowledge structures or routines are broken down (after which they are reassembled), and can thus be characterized by a short peak in Entropy (see also Stephen et al., 2009 and Stephen & Dixon, 2009).

consecutive Entropy values formed the data for subsequent Latent Class Growth Analysis. This analysis was used to identify groups based on the form of the Entropy-trajectories, that is, to distinguish between conversations that could be characterized by a higher Entropy-level followed by a drop in Entropy (i.e. peak-Entropy) and conversations that did not show this pattern of reorganization. Moreover, improvement of children's externalizing behaviour problems was independently assessed through pre- and post-treatment clinicians' ratings. Based on criteria for clinically significant improvement, these ratings were also used to divide the sample into classes: improvers and non-improvers. Consequently, the two estimates of class membership were compared. The results showed that dyads in the peak-Entropy-class belonged more frequently to the improvers-class. To assess whether this finding could be simply attributed to either a decline in frequency of negative dyadic affective states or an increase in positive dyadic affective states, the researchers additionally calculated the Recurrence Rates of each coding category of each conversation (again, 95% confidence intervals were based on 15,000 bootstrap replications). The results from a non-parametric test (Kolmogorov Smirnov test) applied to these not normally distributed data showed no differences between classes in the level of recurrence of any of the affective state categories. This indicates that it might be necessary for people to have a period of unpredictability and flux, in which they try out and explore new behaviours, to develop.

11.2.4   Present Study

To reiterate, in this study, we are interested in teacher learning through reflection in the workplace. Building on a situated and dynamic perspective, learning experiences can be seen as emerging from acting upon information in the (social) environment after a period of time. Through reflection on their working environment, teachers make information explicit. Through reflection on learning experiences, teachers make new insights (developed or adapted meanings, knowledge, and skills) explicit. By making these things explicit, teachers can share them with colleagues, put them to focussed use, and set priorities concerning what to attend to and how to act in which situation. Moreover, attending to information can occur more frequently than having new insights, and therefore reflection on the working environment can occur more frequently than reflection on learning experiences. As an example of how to investigate teacher learning through reflection as an everyday and ongoing process, we designed a study to explore the routine with which teachers engage in making information explicit, and how that, in comparison to the overall levels thereof, relates to making new insights explicit. The routine of reflecting pertains to the temporal stability of that activity, and thus its dynamics should be assessed. This requires the collection of dense time-series from individual teachers.

Our measurement instruments, measurement intervals, and analytic measures were chosen in correspondence with this conceptualization. In accord with the different expected rates of change, we chose to use daily logs to measure reflection on

the environment and monthly logs to measure reflection on learning experiences. We will explore whether these measurement instruments and measurement intervals are useful for the assessment of the dynamics of learning through reflection (see also Kugler, Shaw, Vincente, & Kinsella-Shaw, 1990).

We used the responses to the daily logs to generate time-series for each participant. Each point in these time-series represents the intensity of reflection on the environment, i.e. the number of reflection moments during a day. The analysis measures for the routine of reflection on the environment were calculated by applying a categorical auto-RQA to each time-series. Recurrence Rate was used as a raw measure of routine and informs about the overall regularity of the reflection process. Determinism was used as a measure of the persistence thereof. The analysis measures for the overall level of reflection on the environment and learning experiences were calculated by simply summing up all responses to the daily and monthly logs, respectively. To investigate the extent to which the overall level and the routine of the intensity of making information explicit co-occurs with the overall intensity of making insights explicit, we generated and inspected scatterplots.

11.3   Method

We used a longitudinal, mixed-method design with convenience sampling to assess the relation between the level and routine of teachers' engagement in reflection on their environments to make information explicit and the level of reflection on learning experiences to make insights explicit. To do so, we asked teachers to fill in daily and monthly logs, including open questions about the salient information they attended to and the learning experiences they had, respectively, for a period of 5 months. Analyses were applied to the time-series of frequencies of filled in log entries.

11.3.1   Sample

This study was conducted in one VET college in the Netherlands in 2011 (see also Oude Groote Beverborg et al., 2015a). Team leaders were asked whether team members were willing to participate in this study, and participation was voluntary. A total of 20 teachers participated. The data from 1 teacher were excluded from the analysis, because the teacher had moved to a different employer (a college offering professional education), and the data from 2 other teachers were excluded, because they started 2 months late. Thus, the effective sample size was 17. The participants were employed in departments that taught law, business administration, ICT, laboratory technology, and engineering to students and that coached other teachers. Thirteen participants were female, and 4 were male. Working days per week ranged from 2 to 5. In order to generate enough data for a substantive time-series, but as a

trade-off between practicality and rigor, the study ran for 5 months: from February until June. During this period, all participants had a 2 weeks' holiday. One participant (P12) stopped participating after 2 months, and another participant (P10) after 3 months.

11.3.2 Measurement

The study consisted of two logs: a daily and a monthly log. The daily log (diary) asked teachers to make salient information explicit, and thus measured their engagement in reflection on the environment. The monthly log asked teachers to make their insights explicit, and thus measured their reflections on learning experiences. The logs were designed as short, structured interviews with a few open questions. Thereby, participants could report the information that was most relevant to them individually at each measurement occasion. More specifically, the diaries asked about the most salient information that day and the context, in which the information was attended to. The diary questions were focussed on information from colleagues (de Groot, Endedijk, Jaarsma, Simons, & van Breukelen, 2014). The main diary question was: "What did your colleague say or do that was most salient today?" It was made explicit that this could be something someone said, someone did, something that was read, and so on. Other open questions related to the task the participants worked on for which the reported information was relevant, and to how they responded to the information (see Appendix A for the complete specification of one diary entry translated into English). The diaries were designed in such a way that teachers could report their own experiences. The diaries were therefore sensitive to local and personal circumstances and measured with such a density that fluctuations could be expected to be measurable. The monthly logs were designed similarly and asked to report the learning experiences participants had had sometime in the last month as accurately as possible (Endedijk, 2010). The most important question was: "What have you learnt in the last month?" Additionally, questions about the context the learning experience came from, or in which context it had to be understood, were asked, such as about the task and the goal they related to, what means helped to learn it, the manner in which it was learnt, and in what manner participants realized they had learnt something. Lastly, the monthly log also asked questions about what teachers were satisfied with in their learning process and what could be improved in the future, what goals they would pursue in the future, and what they would attend to in the future (see Appendix B for the full specification of one monthly log entry translated into English).

Diaries were administered on each person's working days and monthly logs on the first working day of the new month. In order to constrain the burden of repeatedly filling in logs, a maximum of three diary entries (making information explicit) and three monthly log entries (making insights explicit) could be filled in per measurement occasion. Also, participants were instructed to spend no more than 5 min on each diary entry (thus a maximum of 15 min per day), and no more than 10 min

on each monthly log entry (thus a maximum of 30 min per month). Teachers were asked to fill in at least one log entry per measurement occasion, but this was not mandatory. Logs were administered online. For each participant's measurement occasion's log, an invitation was sent by email. On some measurement occasions, some invitations failed to be sent. See Fig. 11.2 and Table 11.1 for frequencies of reporting and descriptives. The analyses were applied to the time-series of frequencies of filled in log entries.

In order to uphold motivation, the first author offered individual coaching sessions to the participants. These sessions took place once every month, lasted about 45 min, and were conducted over the telephone. In general, during a session, the information a participant had reported in the log was summarized, and the participant was asked to respond to that. Towards the end of the conversation, the first author categorized some of the information in the diaries and labelled this summary, after which there was opportunity for the participant to reflect upon the labelling of the information. Each conversation ended with the first author asking feedback on the instrument and the conversation. These calls were not intended as part of the measurement of the study and have therefore not been recorded.

11.3.3 Analysis Strategy

The aim of the analyses was to assess in which way the overall level and the routine of the intensity of making information explicit relates to the overall intensity of making insights explicit. We calculated one measure for making insights explicit: each participant's mean of moments of reflection on learning experiences over the measurement period per month participated (overall insight intensity). This measure is based on the monthly log data. The mean per month was calculated to correct for differences between participants in the duration that they participated.

Crucially, this measure was also used to assess whether participants had affinity for the measurement instruments, that is, whether teachers disengaged from the immediacy of their work to make time to 'interact' with our measurement instruments. In line with our request to fill in at least one log entry per measurement occasion, we set a mean of 1 or more reflections on learning experiences per month as the criterion of affinity. Using the monthly log data to categorize participants into groups thus allowed us to differentiate between participants with regard to the validity of administering logs to them. Moreover, it allowed us to contrast group patterns of dynamics of reflection on the environment, which helps to interpret the results.

For reflection on the environment, we calculated three measures. These measures were based on the daily log data. The first measure was the mean of the intensity of making information explicit in the measurement period per working day (overall information intensity). The mean per working day was calculated to correct for differences between participants in working days.

To assess teachers' routine (or within-person variability) in making information explicit, we applied categorical RQA on each participant's time-series of intensities
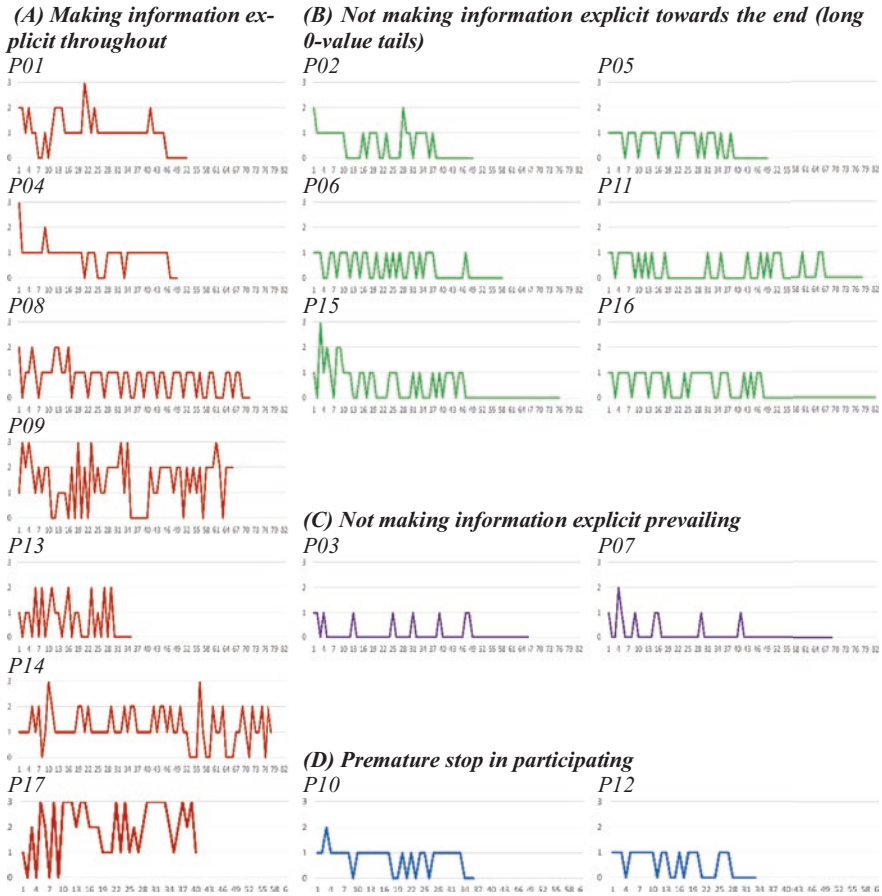
Fig. 11.2 Time-series of participants' intensities of reflection on the environment

Note: Reflection intensity = the number of reflection moments per working day. P stands for participant. Numbers indicate the participants. For each graph, time is on the x-axis and reflection intensity is on the y-axis. The time-series only include those days, on which participants received invitations to fill out daily logs (working days). Consequently, the time-series vary in length. The largest number of working days of a participant during the measurement period was 82 and, to ease comparison, this value was set as the length of each x-axis. The time-series have been categorized based on the participants' response patterns. (**a**): Mean amount of reflection on learning experiences per month is greater than or equal to 1 ($m_{insights} \geq 1$); (**b–d**): Mean amount of learning experiences per month is less than 1 ($m_{insights} < 1$). The participants categorized in (**a**) made information explicit using the measurement instrument throughout the measurement period. The participants categorized in (**b**) did not make information explicit using the measurement instrument towards the end of the measurement period (time-series with long 0-value tails), those in (**c**) had time-series in which 0 (no information made explicit using the measurement instrument on a day) prevailed, and those in (**d**) stopped participating prematurely. Consequently, the participants categorized in (**a**) are considered to have more affinity for our measurement instruments, whereas the participants categorized in (**b**, **c** and **d**) are considered to have less affinity for them. See Table 11.1 for participants' measures in each group

Table 11.1 Descriptives and measures of each participant

Participants	Participation capacity		Daily log measures					Monthly log measures		
	FTE	t_{weeks}	t_{days}	Σ_{infos}	m_{infos}	%REC	%DET	t_{months}	$\Sigma_{insights}$	$m_{insights}$
(a) Making information explicit throughout										
01	0.6	18	52(4)	53	1.02	44	76	5(0)	5	1.00
04	0.6	18	49(7)	44	0.90	65	95	5(0)	9	1.80
08	0.8	18	71(3)	54	0.76	47	71	5(0)	7	1.40
09	0.8	18	66(3)	94	1.42	30	47	5(0)	8	1.60
13	0.4	18	35(1)	24	0.69	37	50	5(0)	8	1.60
14	1.0	17	78(4)	92	1.18	40	67	5(0)	11	2.20
17	0.4	18	40(3)	79	1.98	28	42	5(0)	12	2.40
(b) Not making information explicit towards the end (long 0-value tails)										
02	0.6	18	49(5)	25	0.51	46	78	5(0)	0	0.00
05	0.6	18	49(7)	29	0.59	51	74	5(0)	3	0.60
06	0.8	18	58(11)	23	0.40	51	73	5(0)	3	0.60
11	1.0	18	78(13)	22	0.28	59	81	5(0)	0	0.00
15	1.0	17	76(5)	31	0.41	51	75	4(1)	0	0.00
16	1.0	18	82(5)	30	0.37	53	78	5(0)	2	0.40
(c) Not making information explicit prevailing										
03	0.8	18	66(6)	9	0.14	76	94	5(0)	0	0.00
07	0.8	18	69(4)	9	0.13	79	98	5(0)	2	0.40
(d) Premature stop in participating										
10	1.0	11	36(13)	28	0.78	57	80	3(0)	2	0.67
12	1.0	7	33(2)	19	0.58	49	76	n(0)	0	0.00

Note: FTE = Full-Time Equivalent. Here it stands for the number of days per week a participant is employed by the VET college. 1.0 represents an employment of 5 days per week. t_{weeks} = the number of weeks that the participants participated. The measurement period was 18 weeks. Two participants started 1 week later and 2 participants stopped prematurely. t_{days} = the number of working days (i.e. days on which daily log invitations were sent). The value between parentheses is the amount of invitations whose sending had failed. Σ_{infos} = the overall intensity of making information explicit (i.e. the total number of moments of reflection on the environment in the period). Participants could fill in a maximum of 3 daily log entries per working day. m_{infos} = the mean intensity of making information explicit per working day. This measure was calculated to correct for differences between participants in working days and the duration that they participated. %REC = the Recurrence Rate of daily intensities of making information explicit (i.e. recurrences of the number of reflection moments per working day) during the measurement period (as a percentage). %DET = the Determinism of daily intensities of making information explicit (i.e. the number of reflection moments per working day that recur periodically) in the measurement period (as a percentage). t_{months} = the number of months on which monthly log invitations were sent. The value between parentheses is the amount of invitations whose sending had failed. $\Sigma_{insights}$ = overall intensity of making insights explicit (i.e, the total number of moments of reflection on learning experiences in the period). Participants could fill in a maximum of 3 monthly log entries per month (maximum is 15). $m_{insights}$ = the mean intensity of making insights explicit per month. This measure was calculated to correct for differences between participants in the duration that they participated. The descriptives of the participants have been categorized by their response-patterns. (**a**): $m_{insights} \geq 1$; (**b, c** and **d**): $m_{insights} < 1$. Additionally, the participants categorized in (**a**) made information explicit using the measurement instrument throughout the measurement period. The participants categorized in (**b**) did not make information explicit using the measurement instrument towards the end of the measurement period (time-series with long 0-value tails), those in (**c**) had time-series in which 0 (no information made explicit using the measurement instrument on a day) prevailed, and those in (**d**) stopped participating prematurely. See Fig. 11.2 for graphical representations of the participants' time-series

of reflections on the environment per day. The time-series only include those days, on which participants received invitations to fill in daily logs (working days). Other days, such as weekends or holidays, or days of the week on which participants were not employed or were employed by another employer, are not part of the time-series. These 'non-working days' were cut out to create an uninterrupted time-series. Consequently, the time-series vary in length. The categorical RQA was conducted in MATLAB, using Marwan's toolbox (Marwan et al., 2007; Marwan, Wessel, Meyerfeldt, Schirdewan, & Kurths, 2002). As measures of routine, we used Recurrence Rate as a measure of the overall regularity of the intensity of the reflection process over time, and Determinism as a measure of teachers' persistence in sequences of intensities of reflection. The relations between these four variables were established through visual inspection of scatterplots.

11.4   Results

First, we calculated each measure for each participant. To give an idea of how the trajectories of the intensity of making information explicit (information intensity) correspond to their auto-recurrence plots and their measures, four examples thereof are given in Fig. 11.3.

Second, we assessed the participants' affinity for the measurement instruments. Seven participants had an overall insight intensity (mean of reflections on learning experiences per month) that was greater than or equal to 1, and thus showed more affinity for the measurement instruments. The other ten participants had an overall insight intensity that was less than 1, and thus showed less affinity for the measurement instruments. Splitting the sample into two groups based on overall insight intensity uncovered striking differences in the temporal patterns of making information explicit. Consider Fig. 11.2. The participants categorized in (**a**) made information explicit using the measurement instrument throughout the measurement period, whereas that seems to falter or cease with the participants in (**b**, **c**, and **d**). The participants categorized in (**b**) did not make information explicit using the measurement instrument towards the end of the measurement period (time-series with long 0-value tails), those in (**c**) had time-series in which 0 (no information made explicit using the measurement instrument on a day) prevailed, and those in (**d**) stopped participating prematurely. Consequently, the participants categorized in (**a**) are considered to have, for whatever reason, more affinity for our measurement instruments in the measurement period, whereas the participants categorized in (**b**, **c**, and **d**) are considered to have less affinity for them. Due to the difference between the groups in the fit of the measurement instruments to the participants, administering daily and monthly logs seems to be more valid for the participants in (**a**) than for the others. See Table 11.1 for the participants' measures and descriptives in each group. A comparison of the descriptives of the two groups suggests a connection between affinity for the measurement instruments and the amount of working days and/or the amount of invitations that failed to be sent.
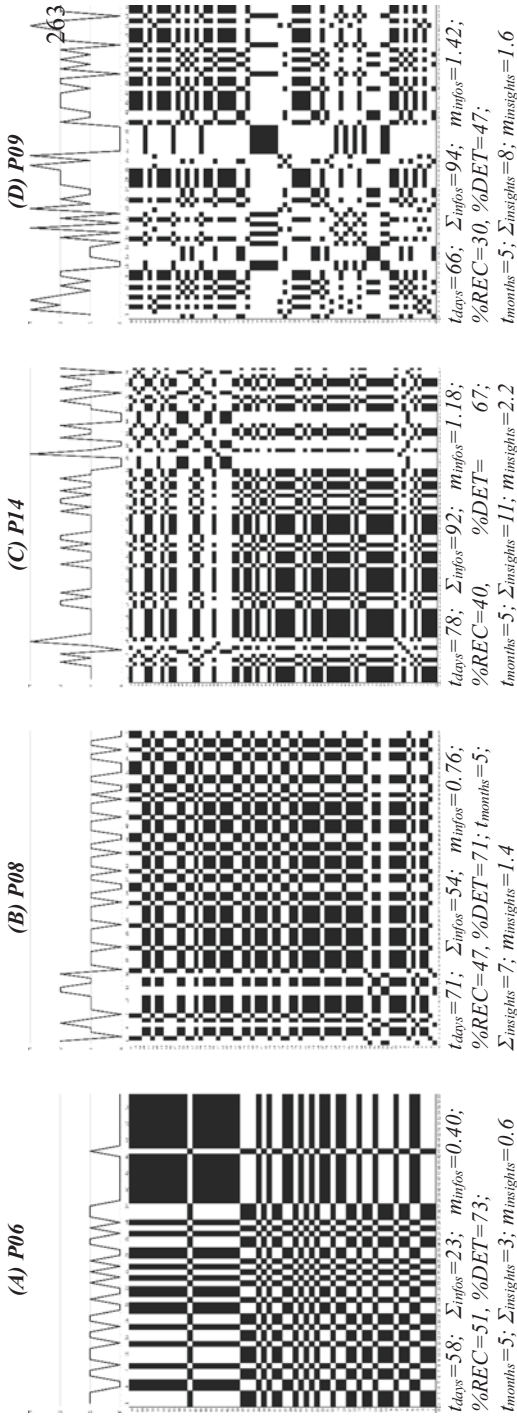
(A) P06

$t_{days}=58$; $\Sigma_{infos}=23$; $m_{infos}=0.40$; $\%REC=51$, $\%DET=73$; $t_{months}=5$; $\Sigma_{insights}=3$; $m_{insights}=0.6$

(B) P08

$t_{days}=71$; $\Sigma_{infos}=54$; $m_{infos}=0.76$; $\%REC=47$, $\%DET=71$; $t_{months}=5$; $\Sigma_{insights}=7$; $m_{insights}=1.4$

(C) P14

$t_{days}=78$; $\Sigma_{infos}=92$; $m_{infos}=1.18$; $\%REC=40$, $\%DET= 67$; $t_{months}=5$; $\Sigma_{insights}=11$; $m_{insights}=2.2$

(D) P09

$t_{days}=66$; $\Sigma_{infos}=94$; $m_{infos}=1.42$; $\%REC=30$, $\%DET=47$; $t_{months}=5$; $\Sigma_{insights}=8$; $m_{insights}=1.6$

Fig. 11.3 Four examples trajectories of intensities of reflection on the environment represented by time-series and recurrence plots

Note: (**a**, **b**, **c** and **d**) are the trajectories of intensities of reflection on the environment of 4 participants (based on daily log data). Each trajectory is represented by a time-series graph and a recurrence plot (top and bottom image, respectively). For descriptions of the time-series, see Fig. 11.2 and Table 11.1. t_{days} = the amount of working days during the measurement period. Σ_{infos} = the sum of intensity of making information explicit in the measurement period. m_{infos} = the mean intensity of making information explicit per working day (overall information intensity). $\%REC$ = the Recurrence Rate of intensities of making information explicit (as a percentage). $\%DET$ = the Determinism of intensities of making information explicit (as a percentage). For detailed descriptions of the the Recurrence Rate and Determinism, see Fig. 11.1 and the Recurrence Quantification Analysis-section. Additionally, the three measures that are based on the monthly logs are also presented: t_{months} = the number of months on which monthly log invitations were sent. $\Sigma_{insights}$ = sum of intensity of making insights explicit in the measurement period. $m_{insights}$ = the mean amount of making insights explicit per month (overall insight intensity). For the extent to which overall insights intensity, overall information intensity, Recurrence Rate, and Determinism correlate, see Fig. 11.4 and the Results-section

Third, we explored how overall insight intensity related to the overall information intensity (mean of reflections on the environment per day), and how both related to the Recurrence Rate of information intensity, and Determinism of information intensity. Consider Fig. 11.4. Figure 11.4 Plot (**a**) suggests a positive correlation between overall information intensity and overall insight intensity within the whole sample, and also within each affinity group separately. More moments of making information explicit co-occurred with more moments of making insights explicit.

Figure 11.4 Plot (**b**) suggests a negative correlation between overall information intensity and Recurrence Rate within the sample, and also within each affinity group separately. More moments of making information explicit co-occurred with less regularity in doing that. This relation might be explained by the increasing difficulty of having an additional reflection moment beyond the previous one on any given day. Note that none of the participants had both a high level of overall information
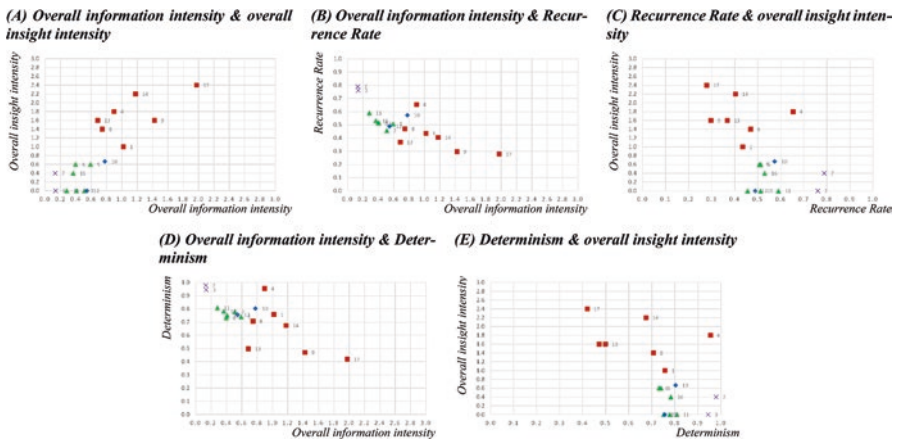
Fig. 11.4 Scatterplots with correlations between overall insight intensity, overall information intensity, Recurrence Rate, and Determinism

Note: Squares represent the group of participants that had more affinity for the measurement instruments (see Fig. 11.2 and Table 11.1). Diamonds, triangles, and crosses represent the group of participants that had less affinity for the measurement instruments. Diamonds represent participants that stopped participating prematurely. Triangles represent participants that did not make information explicit using the measurement instrument towards the end of the measurement period (time-series with long 0-value tails). Crosses represent participants, in whose time-series 0 (no information made explicit using the measurement instrument on a day) prevailed. Numbers indicate the participants. Overall insight intensity = the mean amount of making insights explicit (reflection on learning experiences) per month, overall information intensity = the mean amount of making information explicit (reflection on the environment) per day, Recurrence Rate = Recurrence Rate of information intensities, Determinism = Determinism of information intensities. The means of overall insight intensity (per month) and overall information intensity (per day) for each participant are used to correct for differences between participants in working days and the duration that participants participated. As such, the axis-scales of these two variables go from the minimum (0) to the maximum (3) per measurement occasion. See the text of the Results-section for descriptions of the correlations

intensity and a high Recurrence Rate: a highly regular high level of information intensity did not occur

Figure 11.4 Plot (**c**) suggests a negative correlation between Recurrence Rate and overall insight intensity in the sample. However, within each affinity group separately, there is no clear relation between Recurrence Rate and overall insight intensity. The group of participants that had more affinity for the measurement instruments made more insights explicit and had less regularity in information intensity during the measurement period than the group of participants that had less affinity for the measurement instruments. The level of making insights explicit seems unrelated to the level of regularity of making information explicit when taking affinity for the measurement instruments into account

Figure 11.4 Plot (**d**) suggests a negative correlation between overall information intensity and Determinism within the sample, and also within the group of participants that showed more affinity for the measurement instruments. However, within the group of participants that had less affinity for the measurement instruments, there is no clear relation between overall information intensity and Determinism. Note that in this group, nearly all of the information intensity values were 0 or 1 and that both a re-occurrence of 0 as of 1 creates a recurrence point. Due to this small set of low values, this group of participants had a low level of overall information intensity and a high level of Determinism, which was similar for those participants, whose time-series consisted of more 0's as for those, whose time-series consisted of more 1's. For the group of participants that had more affinity for the measurement instruments, more moments of making information explicit co-occurred with less persistent (periodically recurring) engagement in any of the levels of intensity of making information explicit (or sequences thereof). However, this relation can be explained by the difficulty of maintaining a high level of information intensity over time. Indeed, a highly persistent high level of information intensity did not occur. The correlations from both groups thus highlight the weaknesses of using the response rates of daily logs with several entries as a measurement instrument for the application of RQA

Figure 11.4 Plot (**e**) suggests a negative correlation between Determinism and overall insight intensity in the sample, and also within the group of participants that had more affinity for the measurement instruments. For this group, more moments of making insights explicit co-occurred with less persistent engagement in any of the levels of intensity of making information explicit (or sequences thereof). However, within the group of participants that had less affinity for the measurement instrument, there is no clear relation between Determinism and overall insight intensity. For this group, the level of making insights explicit seems unrelated to the level of persistence of engagement in any of the levels of intensity of making information explicit (or sequences thereof). Following the argumentation given for the relations in plot (**d**), it seems likely that those participants that manage to make information explicit whenever an opportunity occurs are also the ones that are able to make the most insights explicit. Note that, on the one hand, P04 seems to have organized these opportunities as one per day and thereby to be able to make insights explicit, as inferred from a highly persistent moderate level of information intensity

as well as a high level of overall insight intensity. On the other hand, P17 seems to have strived to have as many of these opportunities as possible on each day and thereby to be able to make insights explicit, as inferred from a lowly persistent high level of information intensity as well as a high level of overall insight intensity

In sum, these results point towards a trend that higher levels of overall information intensity and overall insight intensity concur within a certain period of time. On top of that, no clear pattern was found relating the level of overall insight intensity to with which routine participants made information explicit during a certain period of time.

11.5   Discussion

To summarize, in this study, we explored teacher learning through reflection as a situated and dynamic process using logs as the measurement instruments and RQA as the analysis technique. More specifically, the study focussed on the routine with which teachers engage in making information explicit (reflection on the working environment), and how that, in comparison to the overall levels thereof, relates to making new insights explicit (reflection on learning experiences). We also explored the validity of the measurement instruments and measurement intervals for the application of RQA. Seventeen VET teachers filled in daily and monthly logs over a period of 5 months. From the responses to the daily logs, we generated time-series of the intensity of making information explicit (information intensity) for each participant and applied categorical auto-RQA to each time-series. As measures of the routine of information intensity, Recurrence Rate (regularity) and Determinism (persistence) were used. In addition, we calculated a measure for overall information intensity (the mean amount of information intensity per day in the measurement period) and a measure for overall insight intensity (the mean amount of making insights explicit per month in the measurement period). Relations between the four variables were established through inspection of scatterplots. We found that the sample could be divided into two groups: One that had more and one that had less affinity for the measurement instruments. Moreover, inspection of the scatterplots indicated that higher levels of overall information intensity related to higher levels of overall insight intensity. However, the regularity and the persistence of the intensity with which participants made information explicit had no clear relation with the level of overall insight intensity when taking affinity for the measurement instruments into consideration. In this section we will elaborate on these findings.

That the sample could be divided into one group that had more and another group that had less affinity for the measurement instruments (both daily and monthly logs), may be due to several related reasons. One reason might be related to the difference between the groups in the amount of invitations that failed to be sent. The participants in the less affinity group did not receive an invitation about twice as often as the participants in the more affinity group when correcting for the amount of working days. Increasingly, undependability may have led teachers to falter or

cease using our measurement instruments. One of the challenges in conducting this study was to send personalized logs with personalized intervals using an online instrument that was not designed for that, but rather for large-scale, cross-sectional surveys. The developments in digital technology, such as smartphone applications, will have made this problem obsolete for future studies, however.

A second reason might be related to the difference between the groups in the number of days per week they worked. The participants in the less affinity group worked roughly a day more than the participants in the more affinity group, and may simply have been too busy to disengage from the immediacy of their work to make time to reflect by using logs.

A third reason might be related to the dynamics of the reflection process itself. As experience grows, people become less responsive to new information in their environment, and the new information is not further corroborated into experience (Schöner & Dineva, 2007). In this study, the daily logs served as impulses to become aware of information in the environment that some participants might otherwise not have made explicit. Consequently, as experience with this initially attended-to information grew, participants may have felt a need to consolidate acting upon that information first, rather than attending to even more information and deciding how to act upon that. This reason seems particularly fitting for the participants, who did not make information explicit using the measurement instrument towards the end of the measurement period. Nevertheless, whereas administering logs seems to be less valid for these particular participants, the dense time-series the logs generated did point towards an interesting dynamic that future research may explore further.

This third reason relates to that teachers need time to learn (and can attend to teaching less), and also need time to teach (and can attend to learning less) (Mulford, 2010), which points towards the fourth reason: Despite the fact that all teachers volunteered to participate, it could have been that the participants in the more affinity group had a period in which they could attend to learning more, whereas the participants in the less affinity group had a period in which they had to attend to teaching. This fourth reason might complement the second reason.

One final reason may be that the participants did develop and adapt their teaching practices, but not through reflection on the working environment and learning experiences at a later point. Rather, they may have engaged in experimentation with new teaching methods or keeping up to date with the latest literature (Oude Groote Beverborg, Sleegers, & van Veen, 2015c). Despite their initial willingness to participate, they may have found that making information and insights explicit by using logs did not befit them. Future studies could investigate for whom what knowledge content is discovered with what additional learning activities or other forms of reflection. All in all, using daily and monthly logs with open questions to study learning through reflection fitted better to some participants than to others.

For the discussion about the findings related to how the extent to which the overall level and the routine of the intensity of making information explicit co-occurs with the overall intensity of making insights explicit, we focus on the group of participants that was considered to have more affinity for the measurement instruments. We found that levels of overall reflection on the working environment

positively related to levels of overall reflection on learning experiences. In this regard, it is relevant that information to be made explicit is always present in the working environment. Insights, on the contrary, can only be made explicit when learning experiences occurred. As such, the situated manner in which we assessed teacher learning through reflection corroborates findings from large-scale survey studies, which showed that engaging in learning activities more goes together with having more learning results (Oude Groote Beverborg et al., 2015a; Sleegers et al., 2014).

Furthermore, we found no clear relation between the measures of the routine with which teachers reflect on the working environment and their overall reflection on learning experiences. The regularity of making information explicit was unrelated to the overall level of making insights explicit. The persistence of making information explicit could be seen as negatively correlated to the overall level of making insights explicit, but the dispersion was high. To illustrate, from the top three participants in making insights explicit, one had the least and one had the most persistence in the intensity of making information explicit. Thus, the answer to the question about whether learning can be facilitated through reflecting very constantly or in bursts, is: both. The application of RQA thereby extends research on sequences of (multiple) learning activities (Endedijk, Hoekman, & Sleegers, 2014; Zwart et al., 2008). Moreover, these RQA-based findings suggest that constancy in reflection-intensity is not necessarily beneficial to school improvement and educational change (see also Mulford, 2010; Weick, 1996). Such constancy may, again, fit better to some than to others. Consequently, teachers cannot be discharged from the responsibility of finding out what manner of learning befits them personally, and colleagues can only seduce them to do so. Studies with additional measures and in additional contexts are needed to validate our findings concerning the constancy of everyday teacher learning.

How, then, to support teachers in sustaining levels of reflection without enforcing high constancy thereof (see also Giles & Hargreaves, 2006; Timperley & Alton-Lee, 2008)? An answer thereon may not be based on focussing on the routine of engagement in the learning activity itself, but by also taking the situated nature of the process in consideration (Barab et al., 1999). Our findings suggest that those participants that are able to make the most insights explicit are also the ones that manage to make information explicit whenever an opportunity occurs, which could be done by organizing such opportunities (i.e. moments of disengagement from the work flow, the use of evaluation instruments or logs, classroom observations, meetings, or appraisal interviews) with determined intervals, but also by being keen to have as many such moments as the working environment may provide each day, or a combination of both. Either way, the working environment would have to provide ample information that is salient and interesting enough to further think about and to distil a new way of acting from, whenever teachers have an opportunity to do so (Lohman & Woolf, 2001). In this respect, critically reflecting colleagues and transformational school leaders, who inspire, support, and stimulate, are crucial in helping to see the workplace in a new light and in providing examples of how one can synchronize one's practice with newly found information (Hoekstra & Korthagen, 2011; Oude

Groote Beverborg et al., 2015c; van Woerkom, 2010). Future research could investigate the development and dynamics of coordination of team members in creating such an interesting environment by engaging in knowledge sharing with the aim to co-construct shared meaning and to facilitate school improvement and educational change (see also Zoethout, Wesselink, Runhaar, & Mulder, 2017).

In sum, the findings of this study indicate that teachers who make more information from their working environment explicit are also able to make more new insights explicit. This suggests that higher levels of engagement in reflection are beneficial to teachers' developments, and, by extension, to educational change and school improvement. The routine with which teachers make information explicit was found to be mostly unrelated to making new insights explicit. Of importance seems to be to reflect upon the working environment whenever an opportunity arises. Crucial seems to be that this (social) environment provides information that is salient and interesting enough to distil a new way of acting and attending from. Teachers might additionally benefit sometimes from organizing opportunities to become aware of information in the environment with a certain constancy. In this regard, the use of daily and monthly logs seems to fit better to some participants than to others.

This study is a first step in understanding teacher learning through reflection in the workplace as an everyday and ongoing process. The use of measurement instruments that generate dense time-series and the application of RQA to assess stability and flexibility over time shows that longitudinal research can concentrate on more than just on growth or couplings between variables over time (e.g. Hallinger & Heck, 2011; Heck & Hallinger, 2009; Heck & Hallinger, 2010; Oude Groote Beverborg et al., 2015a; Sleegers et al., 2014; Smylie & Wenzel, 2003; Thoonen, Sleegers, Oort, Peetsma, & Geijsel, 2011). Moreover, the study provides an example of how novel methodology, such as RQA, can be adopted to tap into professional learning as a dynamic and situated process in support of school improvement and educational change.

11.5.1 Limitations & Future Directions

The initial idea of the study was to dive deeper into the reflection process than presented here, by measuring what specific types of information teachers attended to using the daily logs, by measuring the contents of learning experiences using the monthly logs, by analysing the dynamics of attending to those types of information using categorical auto-RQA, and by establishing a relation between for instance the persistence in one type of information and the occurrence of a learning experience with a corresponding content. With this aim, we coded the daily and monthly log entries. However, the time-series generated per code-category were not dense enough for the application of RQA. Moreover, we assumed that setting a fixed time for reporting learning experiences would help generate a higher response rate. However, not knowing when learning experiences took place during the months

made it very difficult to relate it to the information reported in the daily logs. Thus, the design failed to generate the timing information that would have been needed to be able to model the learning experiences' occurrences. Having participants fill in learning experiences at (or very soon after) the moment they have them, would therefore have been a better approach. Additionally, our choice of measurement interval was a compromise between the expected rate of change with which salient information would be made explicit and the practical consideration of not wanting to burden the participating teachers too much. Our measurement intervals were therefore too crude for our initial purposes. In sum, measurement methods with a higher sampling rate, such as observations that happen in real-time, are needed to model how information in the working environment affords development and adaptation more accurately (Granic & Patterson, 2006; Lewis et al., 1999; Lichtwarck-Aschoff et al., 2012). Nevertheless, qualitative analyses on the data generated by the logs used in this study can be used to relate the contents that teachers reflected upon with the contents of what they learnt. This would still contribute to understanding more about the role of affordances in teacher learning, but the aim would no longer lie on finding systematic relationships (Barab & Roth, 2006; Greeno, 1994; Little, 2003; Maitlis, 2005).

We would like to stress that RQA's derive their power from frequent measurements – and not from a large sample size. Whereas using small samples could constrain generalizability, studies assessing for instance the temporal pattern of teacher interactions in only one team in real-time, might provide important, new insights into the process of how teachers collaborate to make sense of the challenges they face and how that culminates in the generation of new knowledge or a shared meaning (e.g. Fullan, 2007). Additionally, such studies might prove very valuable for researchers, who are interested in the systematics of change processes and seek to combine the results of various studies in simulation studies (Clarke & Hollingsworth, 2002), rather than meta-analyses (see also Richter, Dawson, & West, 2011; Sun & Leithwood, 2012; Witziers, Bosker, & Krüger, 2003). By building on the current study, future research could contribute to a bottom-up understanding of how learning communities, but also change capacity of schools, emerge and continue to evolve (Hopkins, Harris, Stoll, & Mackay, 2010; Stoll, 2009). Another benefit of the proposed measurement methods and analyses, due to their focus on the circumstances and periodicities of individuals, is that they allow for tailored advice to individual teachers (or teams of teachers). Consequently, this approach to investigating professional learning would allow teachers and policy makers alike to formulate situated expectations about the pace of adaptation, the rate of innovations within a certain time, and delays in proficiency. An interesting follow-up question nevertheless concerns the extent to which diaries served as an intervention for fostering reflective learning and thus influenced the learning occurrences accordingly (Hoekstra & Korthagen, 2011). A new study with an experimental design and additional dependent measures would be needed to investigate this (Maag Merki, 2014).

Despite its limitations, this study does provide a first enquiry into studying teacher learning as a situated and dynamic process through the use of logs and RQA. In future research, the methodology could have utility in studying aspects of

the dynamics of teacher learning such as, on an individual level, shifts in appreciation of the importance of certain classroom practices or differentiation in perception, or on an organizational level, alternations of periods of tight versus loose couplings between teachers, teams, or departments (see also Korthagen, 2010; Kunnen & Bosma, 2000; Mulford, 2010; Nonaka, 1994; Orton & Weick, 1990). The methodology could also help policy makers in balancing top-down and bottom-up processes in shaping the organization of the school (e.g. Feldhoff, Huber, & Rolff, 2010; Hopkins et al., 2010; Spillane et al., 2002; van der Vegt & van de Vliert, 2002). Moreover, by studying the temporal pattern of sensemaking processes in schools (see also Coburn, 2001; Feldhoff & Wurster, 2017; Spillane et al., 2002), more can be understood about the development of professional learning communities and the inner workings of the change capacities of schools. Consequently, in line with trends in accountability to focus on learning of organizations rather than fulfilment of inspection criteria, Inspectorates of Education could use the methodology to tap into a developmental process rather than only the results thereof in order to support the sensemaking processes in schools (Feldhoff & Wurster, 2017).

Appendices

Appendix A

Daily Log 1(2)

Information

This question is about informal learning from colleagues in the workplace.

Informal learning can be seen as the daily discovery of information.

Information can be known or new, it can be positive or negative, and it can be something from the educational praxis or something from a conversation.

More concretely, you can think of information as something a colleague said; something that was recommended to you; something you experienced; the manner in which you did something; the feedback you gave someone; something you did not do; etc.

This question is about which information struck you the most today. Below you see four answer categories.

Below, you see four answer categories.

Choose one of the answer categories.

Later, you can choose a new answer category.

After you have clicked on one of the options, you will be presented with questions about the nature of the information that struck you.

After you have answered the questions about the nature of the information, you can choose one of the four answer categories again.

You can choose an answer category maximally three times, thereafter the diary entry of today will stop.

Try to use no more than 5 min for filling in today's diary entries.

Which of the options below struck you the most today?
(Where "colleague" is stated, you can also read "colleagues")

☐ I agreed with something a colleague said or did

☐ I disagreed with something a colleague said or did

☐ Something a colleague did helped me

☐ Something a colleague did hindered me

$\boxed{\textit{PREVIOUS page}}$ $\boxed{\textit{NEXT page}}$

Daily Log 2(2)

Information

Where "colleague" is stated, you can also read "colleagues".

You stated that you agreed with something a colleague said or did today.[6]

The following questions elaborate on that.

Try to answer the open questions in no more than three sentences.

---

[6] In case another answer category was selected on the previous page, the text throughout this page was adapted accordingly.

What did your colleague say or do today?

What about what your colleague said or did was relevant for you?
(If needed, you can select more than one option, but try to constrain your answer to one option.)

☐ That a colleague said or did something
☐ That that colleague said or did something
☐ What the colleague said or did
☐ Something about what the colleague said or did (e.g., that one sentence or action)
☐ The result of what the colleague said or did
☐ All of the colleagues performance
Otherwise, namely…

What was the task that you worked on, to which what your colleague said or did related?

What was your reaction to what your colleague said or did?

To what extent did you agree with what your colleague said or did?

☐ 1: I agreed a little
☐ 2: I agreed
☐ 3: I agreed a lot

Do you intend to attend to it in the following weeks?

☐ Yes
☐ No
☐ Does not apply

PREVIOUS page		*NEXT page*

Appendix B

Monthly Log

Learning Experience

Learning can occur everywhere and always. Learning can be planned and spontaneous. You become conscious of having learned something when you have had a learning experience.

You can think for example of a learning experience as having found a new way to prepare a task with your colleagues, or as having had an insight about how you can transfer something to your students after having had a conversation with a colleague.

The questions in the monthly log are about learning experiences that you have had in the past month. We kindly ask you to report three learning experiences.[7] Each entry is about one learning experience. This is the entry of learning experience 1[8].

Try to answer the questions in no more than three sentences.

---

[7] Although we kindly asked to report three learning experiences, it was voluntary whether participants filled in 0, 1, 2, or 3 monthly log entries.

[8] For the second and third entry filled in within the log of 1 month, this number is 2 or 3, respectively.

1. What did you learn in the past month?

2. For the performance of which task, was what was learned relevant?

3. To which personal or professional development goal did what was learned relate?

4. What was needed to learn it?
(Think for instance of what knowledge, skills, experiences, means, or people)

5. In which way did you learn it?

6. Why do you learn it in this specific way?

7. How did you find out that you had learned something?
Describe the learning experience.
(i.a. with whom, working on which task, etc.)

8. With which aspects of the learning process are you satisfied, and what would you do differently next time?

9. Now that you have learned this, what will you attend to in the following weeks?

10. On the basis of this learning experience, which personal or professional goal do you set for yourself for the following weeks?

PREVIOUS page   *NEXT page*

References

Argyris, C., & Schön, D. A. (1974). *Theory in practice: Increasing professional effectiveness*. San Francisco, CA: Jossey-Bass.

Barab, S. A., Cherkes-Julkowski, M., Swenson, R., Garrett, S., Shaw, R. E., & Young, M. (1999). Principles of self-organization: Learning as participation in autocatakinetic systems. *Journal of the Learning Sciences, 8*(3–4), 349–390.

Barab, S. A., & Roth, W. M. (2006). Curriculum-based ecosystems: Supporting knowing from an ecological perspective. *Educational Researcher, 35*(5), 3–13.

Bolger, N., & Laurenceau, J. P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. New York, NY: Guilford Press.

Brick, T. R., Gray, A. L., & Staples, A. D. (2018). Recurrence quantification for the analysis of coupled processes in aging. *The Journals of Gerontology: Series B, 73*(1), 134–147.

Broer, H., & Takens, F. (2009). *Dynamical systems and chaos*. Utrecht, The Netherlands: Epsilon Uitgaven.

Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education, 18*, 947–967.

Coburn, C. E. (2001). Collective Sensemaking about Reading: How teachers mediate Reading policy in their professional communities. *Educational Evaluation and Policy Analysis, 23*(2), 145–170.

Coburn, C. E. (2004). Beyond decoupling: Rethinking the relationship between the institutional environment and the classroom. *Sociology of Education, 77*, 211–244.

Coburn, C. E. (2005). Shaping teacher sensemaking: School leaders and the enactment of reading policy. *Educational Policy, 19*(3), 476–509.

Coburn, C. E. (2006). Framing the problem of reading instruction: Using frame analysis to uncover the microprocesses of policy implementation. *American Educational Research Journal, 43*(3), 343–349.

Cox, R. F. A., Van der Steen, S., Guevara Guerrero, M., Hoekstra, L., & Van Dijk, M. (2016). Chromatic and anisotropic cross-recurrence quantification analysis of interpersonal behavior. In C. Webber, C. Ioana, & N. Marwan (Eds.), *Recurrence Plots and Their Quantifications: Expanding Horizons: Proceedings of the 6th International Symposium on Recurrence Plots*, Grenoble, France, 17–19 June 2015. (Springer Proceedings in Physics). Springer.

Dale, R., & Spivey, M. J. (2005). Categorical recurrence analysis of child language. In *Proceedings of the 27th annual meeting of the cognitive science society* (pp. 530–535). Mahwah, NJ: Lawrence Erlbaum.

de Groot, E., Endedijk, M. D., Jaarsma, A. D. C., Simons, P. R. J., & van Breukelen, P. (2014). Critically reflective dialogues in learning communities of professionals. *Studies in Continuing Education, 36*(1), 15–37.

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

Endedijk, M. D. (2010). *Student teachers' self-regulated learning*. (Doctoral dissertation). Utrecht, The Netherlands: University Utrecht.

Endedijk, M. D., Brekelmans, M., Verloop, N., Sleegers, P. J. C., & Vermunt, J. D. (2014). Individual differences in student teachers' self-regulated learning: An examination of regulation configurations in relation to conceptions of learning to teach. *Learning and Individual Differences, 30*, 155–162.

Endedijk, M. D., Hoekman, M., & Sleegers, P. J. C. (2014). Learning paths of engineers: Studying sequences of learning activities to understand knowledge workers' professional development. In *7th EARLI SIG 14 Conference.* Oslo, Norway, 27–29 August 2014b.

Feldhoff, T., Huber, S. G., & Rolff, H. G. (2010). Steering groups as designers of school development processes. *Journal for Educational Research Online, 2*(2), 98–124.

Feldhoff, T., Radisch, F., & Bischof, L. M. (2016). Designs and methods in school improvement research: A systematic review. *Journal of Educational Administration, 54*(2), 209–240.

Feldhoff, T., Radisch, F., & Klieme, E. (2014). Methods in longitudinal school improvement: State of the art. *Journal of Educational Administration, 52*(5).

Feldhoff, T., & Wurster, S. (2017). Ein Angebot das sie nicht ablehnen können? Zur Funktion von Schulinspektionsergebnissen als Deutungsangebot zur Unterstützung schulischer Verarbeitungsprozesse und schulische Reaktionsweisen auf das Angebot. Empirische Pädagogik.

Fullan, M. (2007). *The new meaning of educational change* (4th ed.). London, UK: Teachers College Press.

Geursen, J., de Heer, A., Korthagen, F. A., Lunenberg, M., & Zwart, R. (2010). The importance of being aware: Developing professional identities in educators and researchers. *Studying Teacher Education, 6*(3), 291–302.

Gibson, J. J. (1979/1986). *The ecological approach to visual perception*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

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.

Granic, I., & Dishion, T. J. (2003). Deviant talk in adolescent friendships: A step toward measuring a pathogenic attractor process. *Social Development, 12*(3), 314–334.

Granic, I., & Patterson, G. R. (2006). Toward a comprehensive model of antisocial development: A dynamic systems approach. *Psychological Review, 113*(1), 101–131.

Greeno, J. G. (1994). Gibson's affordances. *Psychological Review, 101*(2), 336–342.

Greeno, J. G. (1998). The situativity of knowing, learning, and research. *American Psychologist, 53*(1), 5–26.

Guastello, S. J. (2002). *Managing emergent phenomena. Nonlinear dynamics in work organizations*. Mahwah, NJ: Lawrence Erlbaum Associates.

Guastello, S. J., Johnson, E. A., & Rieke, M. L. (1999). Nonlinear dynamics of motivational flow. *Nonlinear Dynamics, Psychology, and Life Sciences, 3*, 259–273.

Hallinger, P., & Heck, R. (2011). Exploring the journey of school improvement: Classifying and analyzing patterns of change in school improvement processes and learning outcomes. *School Effectiveness and School Improvement, 22*, 149–173.

Hamaker, E. L. (2012). Why researchers should think "within-person": A paradigmatic rationale. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 43–61). New York, NY: Guilford.

Heck, R. H., & Hallinger, P. (2009). Assessing the contribution of distributed leadership to school improvement and growth in math achievement. *American Educational Research Journal, 46*, 626–658.

Heck, R. H., & Hallinger, P. (2010). Collaborative leadership effects on school improvement: Integrating unidirectional- and reciprocal-effects models. *The Elementary School Journal, 111*(2), 226–252.

Heft, H. (2001). *Ecological psychology in context: James Gibson Roger Barker and the legacy of William James's radical empiricism*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Hoekstra, A., & Korthagen, F. (2011). Teacher learning in a context of educational change: Informal learning versus systematically supported learning. *Journal of Teacher Education, 62*(1), 76–92.

Hopkins, D., Harris, A., Stoll, L., & Mackay, T. (2010). School and system improvement: State of the art review. Keynote presentation prepared for the *24th. International Congress of School Effectiveness and School Improvement*.

Horn, I. S. (2005). Learning on the job: A situated account of teacher learning in high school mathematics departments. *Cognition and Instruction, 23*(2), 207–236.

Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.

Korthagen, F., & Vasalos, A. (2005). Levels in reflection: Core reflection as a means to enhance professional growth. *Teachers and Teaching: Theory and Practice, 11*(1), 47–71.

Korthagen, F. A. J. (2001). *Linking practice and theory: The pedagogy of realistic teacher education*. Paper presented at the Annual Meeting of the American Educational Research Association. Seattle, WA.

Korthagen, F. A. J. (2010). Situated learning theory and the pedagogy of teacher education: Towards an integrative view of teacher behaviour and teacher learning. *Teacher and Teacher Education, 26*(1), 98–106.

Kugler, P. N., Shaw, R. E., Vincente, K. J., & Kinsella-Shaw, J. (1990). Inquiry into intentional systems I: Issues in ecological physics. *Psychological Research, 52*(2–3), 98–121.

Kunnen, E. S., & Bosma, H. A. (2000). Development of meaning making: A dynamic systems approach. *New Ideas in Psychology, 18*(1), 57–82.

Lave, J., & Wenger, E. (1991). *Situated learning. Legitimate peripheral participation*. New York, NY: Cambridge University Press.

Lewis, M. D., Lamey, A. V., & Douglas, L. (1999). A new dynamic systems method for the analysis of early socioemotional development. *Developmental Science, 2*(4), 457–475.

Lichtwarck-Aschoff, A., Hasselman, F., Cox, R., Pepler, D., & Granic, I. (2012). A characteristic destabilization profile in parent-child interactions associated with treatment efficacy for aggressive children. *Nonlinear Dynamics, Psychology, and Life Sciences, 16*(3), 353–379.

Lichtwarck-Aschoff, A., Kunnen, S. E., & van Geert, P. L. (2009). Here we go again: A dynamic systems perspective on emotional rigidity across parent–adolescent conflicts. *Developmental Psychology, 45*(5), 1364–1375.

Little, J. (1990). The persistence of privacy: Autonomy and initiative in teachers' professional relations. *The Teachers College Record, 91*(4), 509–536.

Little, J. W. (2003). Inside teacher community: Representations of classroom practice. *Teachers College Record, 105*(6), 913–945.

Little, J. W., & Horn, I. S. (2007). Normalizing' problems of practice: Converting routine conversation into a resource for learning in professional communities. In L. Stoll & K. S. Louis (Eds.), *Professional learning communities: Divergence depth and dilemmas* (pp. 79–92). London, UK: McGraw-Hill Education (UK).

Lohman, M. C., & Woolf, N. H. (2001). Self-initiated learning activities of experienced public school teachers: Methods, sources, and relevant organizational influences. *Teachers and Teaching, 7*, 59–74.

Lunenberg, M., Korthagen, F., & Zwart, R. (2011). Self-study research and the development of teacher educator's professional identities. *European Educational Research Journal, 10*(3), 407–420.

Lunenberg, M., Zwart, R., & Korthagen, F. (2010). Critical issues in supporting self-study. *Teaching and Teacher Education, 26*(6), 1280–1289.

Maag Merki, K. (2014). Conducting intervention studies on school improvement: An analysis of possibilities and constraints based on an intervention study of teacher cooperation. *Journal of Educational Administration, 52*(5), 590–616.

Maag Merki, K., Grob, U., Rechsteiner, B., Rickenbacher, A., & Wullschleger, A. (2021). Regulation activities of teachers in secondary schools. Development of a theoretical framework and exploratory analyses in four secondary schools based on time sampling data. In A. Oude Groote Beverborg, K. Maag Merki, F. Feldhoff, & F. Radisch (Eds.), *Concept and design developments in school improvement research. State of the art longitudinal, multilevel, and mixed methods and their relevance for educational accountability* (pp. 257–301). Dordrecht, The Netherlands: Springer.

Mainhard, M. T., Pennings, H. J., Wubbels, T., & Brekelmans, M. (2012). Mapping control and affiliation in teacher–student interaction with State Space Grids. *Teaching and Teacher Education, 28*(7), 1027–1037.

Maitlis, S. (2005). The social processes of organizational sensemaking. *Academy of Management Journal, 48*(1), 21–49.

Marwan, N., Romano, M. C., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports, 438*(5–6), 237–329.

Marwan, N., Wessel, N., Meyerfeldt, U., Schirdewan, A., & Kurths, J. (2002). Recurrence plot based measures of complexity and its application to heart rate variability data. *Physical Review E, 66*(2), 26702.

Molenaar, P. C., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychological Science, 18*(2), 112–117.

Mulford, B. (2010). Recent developments in the field of educational leadership: The challenge of complexity. In A. Hargreaves, A. Lieberman, M. Fullan, & D. Hopkins (Eds.), *Second international handbook of educational change* (pp. 187–208). Dordrecht, The Netherlands: Springer.

Ng, F. S. D. (2021). Reframing educational leadership research in the 21st century. In A. Oude Groote Beverborg, K. Maag Merki, F. Feldhoff, & F. Radisch (Eds.), *Concept and design developments in school improvement research. State of the art longitudinal, multilevel, and mixed methods and their relevance for educational accountability* (pp. 107–135). Dordrecht, The Netherlands: Springer.

Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization Science, 5*(1), 14–37.

O'Brien, B. A., Wallot, S., Haussmann, A., & Kloos, H. (2014). Using complexity metrics to assess silent reading fluency: A cross-sectional study comparing oral and silent reading. *Scientific Studies of Reading, 18*(4), 235–254.

Orton, J. D., & Weick, K. E. (1990). Loosely coupled systems: A reconceptualization. *Academy of Management Review, 15*(2), 203–223.

Oude Groote Beverborg, A. (2015). *Fostering sustained teacher learning: Co-creating purposeful and empowering workplaces.* (Doctoral dissertation). Enschede, The Netherlands: University of Twente.

Oude Groote Beverborg, A., Sleegers, P. J. C., Endedijk, M. D., & van Veen, K. (2015a). Towards sustaining levels of reflective learning: how do transformational leadership, task interdependence, and self-efficacy shape teacher learning in schools? *Societies, 5*, 187–219.

Oude Groote Beverborg, A., Sleegers, P. J. C., & van Veen, K. (2015b). Fostering teacher learning in VET colleges: Do leadership and teamwork matter? *Teaching and Teacher Education, 48*, 22–33.

Oude Groote Beverborg, A., Sleegers, P.J.C., & van Veen, K. (2015c). Promoting VET teachers' individual and social learning activities: the empowering and purposeful role of transformational leadership, interdependence, and self-efficacy. *Empirical Research in Vocational Education and Training, 7*(5).

Putnam, R. T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational Researcher, 29*(1), 4–15.

Reed, E. S. (1996). *Encountering the world. Toward an ecological psychology*. New York, NY: Oxford University Press.

Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination: Common ground and the coupling of eye movements during dialogue. *Psychological Science, 18*(5), 407–413.

Richardson, M. J., Schmidt, R. C., & Kay, B. A. (2007). Distinguishing the noise and attractor strength of coordinated limb movements using recurrence analysis. *Biological Cybernetics, 96*(1), 59–78.

Richter, A. W., Dawson, J. F., & West, M. A. (2011). The effectiveness of teams in organizations: A meta-analysis. *The International Journal of Human Resource Management, 22*(13), 2749–2769.

Riley, M. A., & Van Orden, G. C. (2005). *Tutorials in contemporary nonlinear methods for the behavioral sciences.* Retrieved March 1, 2005, from http://www.nsf.gov/sbe/bcs/pac/nmbs/nmbs.jsp

Schön, D. (1983). *The reflective practitioner: How professionals think in action*. New York, NY: Basic Books.

Schöner, G., & Dineva, E. (2007). Dynamic instabilities as mechanisms for emergence. *Developmental Science, 10*(1), 69–74.

Shockley, K., Santana, M. V., & Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance, 29*(2), 326.

Sleegers, P. J. C., & Spillane, J. P. (2009). In pursuit of school leadership and management expertise: Introduction to the special issue. *Leadership and Policy in Schools, 8,* 121–127.

Sleegers, P. J. C., Thoonen, E. E., Oort, F. J., & Peetsma, T. T. (2014). Changing classroom practices: The role of school-wide capacity for sustainable improvement. *Journal of Educational Administration, 52*(5), 617–652.

Sleegers, P. J. C., Wassink, H., van Veen, K., & Imants, J. (2009). School leaders' problem framing: A sense-making approach to problem-solving processes of beginning school leaders. *Leadership and Policy in Schools, 8*(2), 152–172.

Smylie, M. A., & Wenzel, S. A. (2003). The Chicago Annenberg challenge: Successes, failures, and lessons for the future. *Final Technical Report of the Chicago Annenberg Research Project.*

Spillane, J. P., & Miele, D. B. (2007). Evidence in practice: A framing of the terrain. In P. A. Moss (Ed.), *Evidence and decision-making: The 106th yearbook of the National Society for the Study of Education, Part I* (pp. 46–73). Malden, MA: Blackwell Publishing.

Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocussing implementation research. *Review of Educational Research, 72*(3), 387–431.

Spillane, J. P., & Zuberi, A. (2021). Designing and piloting a leadership daily practice log: Using logs to study the practice of leadership. In A. Oude Groote Beverborg, K. Maag Merki, F. Feldhoff, & F. Radisch (Eds.), *Concept and design developments in school improvement research. State of the art longitudinal, multilevel, and mixed methods and their relevance for educational accountability* (pp. 155–195). Dordrecht, The Netherlands: Springer.

Staples, D. S., & Webster, J. (2008). Exploring the effects of trust, task interdependence and virtualness on knowledge sharing in teams. *Information Systems Journal, 18*(6), 617–640.

Steenbeek, H. W., & van Geert, P. L. C. (2007). A theory and dynamic model of dyadic interaction: Concerns, appraisals, and contagiousness in a developmental context. *Developmental Review, 27,* 1–40.

Stephen, D. G., & Dixon, J. A. (2009). The self-organization of insight: Entropy and power laws in problem solving. *The Journal of Problem Solving, 2*(1), 72–101.

Stephen, D. G., Dixon, J. A., & Isenhower, R. W. (2009). Dynamics of representational change: Entropy, action, and cognition. *Journal of Experimental Psychology. Human Perception and Performance, 35*(6), 1811–1832.

Stoll, L. (2009). Capacity building for school improvement or creating capacity for learning? A changing landscape. *Journal of Educational Change, 10,* 115–127.

Sun, J., & Leithwood, K. (2012). Transformational school leadership effects on student achievement. *Leadership and Policy in Schools, 11*(4), 418–451.

Takens, F. (1981). Detecting strange attractors in turbulence. In *Dynamical systems and turbulence, Warwick 1980* (pp. 366–381). Heidelberg, Germany: Springer.

Thoonen, E. E. J., Sleegers, P. J. C., Oort, F. J., Peetsma, T. T. D., & Geijsel, F. P. (2011). How to improve teaching practices: The role of teacher motivation, organizational factors, and leadership practices. *Educational Administration Quarterly, 47*(3), 496–536.

Timperley, H., & Alton-Lee, A. (2008). Reframing teacher professional learning: An alternative policy approach to strengthening valued outcomes for diverse learners. *Review of Research in Education, 32*(1), 328–369.

van der Lans, R. M. (2018). On the "association between two things": The case of student surveys and classroom observations of teaching quality. *Educational Assessment, Evaluation and Accountability, 30*(4), 347–366.

van der Lans, R. M., van de Grift, W. J. C. M., & van Veen, K. (2018). Developing an instrument for teacher feedback: Using the Rasch model to explore teachers' development of effective teaching strategies and behaviors. *The Journal of Experimental Education, 86*(2), 247–264.

van der Vegt, G., & van de Vliert, E. (2002). Intragroup interdependence and affectiveness: Review and proposed directions for theory and practice. *Journal of Managerial Psychology, 17*(1), 50–67.

van Geert, P., & Steenbeek, H. (2005). Explaining after by before: Basic aspects of a dynamic systems approach to the study of development. *Developmental Review, 25,* 408–442.

van Geert, P., & Steenbeek, H. (2014). The good, the bad and the ugly? The dynamic interplay between educational practice, policy and research. *Complicity: An International Journal of Complexity and Education, 11*(2), 22–39.

van Woerkom, M. (2004). The concept of critical reflection and its implications for human resource development. *Advances in Developing Human Resources, 6*(2), 178–192.

van Woerkom, M. (2010). Critical reflection as a rationalistic ideal. *Adult Education Quarterly, 60*(4), 339–356.

Voestermans, P., & Verheggen, T. (2007). *Cultuur en lichaam: Een cultuurpsychologisch perspectief op patronen in gedrag*. Oxford, UK: Blackwell.

Voestermans, P., & Verheggen, T. (2013). *Culture as embodiment: The social tuning of behavior*. Chichester, UK: John Wiley Sons.

Weick, K. E. (1996). Enactment and the boundaryless career: Organizing as we work. In M. B. Arthur & D. M. Rousseau (Eds.), *The boundaryless career: A new employment principle for a new organizational era* (pp. 40–57). New York, NY: Oxford University Press Inc.

Weick, K. E. (2006). The role of imagination in the organizing of knowledge. *European Journal of Information Systems, 15*, 446–452.

Weick, K. E. (2011). Organized sensemaking: A commentary on processes of interpretive work. *Human Relations, 65*(1), 141–153.

Wenger, E. (1998). *Communities of practice: Learning meaning and identity*. New York, NY: Cambridge University Press.

Wijnants, M. L., Bosman, A. M., Hasselman, F. W., Cox, R. F., & Van Orden, G. C. (2009). 1/f scaling in movement time changes with practice in precision. *Nonlinear Dynamics, Psychology, and Life Sciences, 13*(1), 75–94.

Wijnants, M. L., Hasselman, F., Cox, R. F. A., Bosman, A. M. T., & Van Orden, G. (2012). An interaction-dominant perspective on reading fluency and dyslexia. *Annals of Dyslexia, 62*(2), 100–119.

Witziers, B., Bosker, R. J., & Krüger, M. L. (2003). Educational leadership and student achievement: The elusive search for an association. *Educational Administration Quarterly, 39*(3), 398–425.

Zoethout, H., Wesselink, R., Runhaar, P., & Mulder, M. (2017). Using Transactivity to understand emergence of team learning. *Small Group Research, 48*(2), 190–214.

Zwart, R. C., Wubbels, T., Bergen, T. C. M., & Bolhuis, S. (2007). Experienced teacher learning within the context of reciprocal peer coaching. *Teachers and Teaching: Theory and Practice, 13*(2), 165–187.

Zwart, R. C., Wubbels, T., Bolhuis, S., & Bergen, T. C. M. (2008). Teacher learning through reciprocal peer coaching: An analysis of activity sequences. *Teaching and Teacher Education, 24*(4), 982–1002.