Chapter 13

Concept and Design Developments in School Improvement Research: General Discussion and Outlook for Further Research

**Tobias Feldhoff, Katharina Maag Merki, Arnoud Oude Groote Beverborg,
and Falk Radisch**

This book aimed to present innovative designs, measurement instruments, and analysis methods by way of illustrative studies. Through these methodology and design developments, the complexity of school improvement in the context of new governance and accountability measures can be better depicted in future research projects. In this concluding chapter, we discuss what strengths the presented methodologies and designs have and to what extent they do better justice to the multilevel, complex, and dynamic nature of school improvement than previous approaches. In addition, we outline some needs for future research in order to gain new perspectives for future studies.

In this discussion we are guided by Feldhoff and Radisch's framework on complexity (see Chap. 2). The chapters in this volume contribute in particular to discussion of the following aspects:

- The longitudinal nature of the school improvement process
- School improvement as a multilevel phenomenon
- Indirect and reciprocal effects
- Variety of meaningful factors

T. Feldhoff (✉)

Johannes Gutenberg University, Mainz, Germany

e-mail: feldhoff@uni-mainz.de

K. Maag Merki

University of Zurich, Zurich, Switzerland

A. Oude Groote Beverborg

Radboud University Nijmegen, Nijmegen, The Netherlands

F. Radisch

University of Rostock, Rostock, Germany

© The Author(s) 2021

A. Oude Groote Beverborg et al. (eds.), *Concept and Design Developments in School Improvement Research*, Accountability and Educational Improvement, https://doi.org/10.1007/978-3-030-69345-9_13

303

13.1 The Longitudinal Nature of the School Improvement Process

Even though school improvement always implies a change (Stoll & Fink, 1996), studying school improvement longitudinally was surprisingly neglected for a long time (Feldhoff, Radisch, & Klieme, 2014). For this reason, it is particularly important that four of the contributions in this volume (Chaps. 9, 10, 11, and 12) examine school improvement processes longitudinally. All of them use logs as a measurement instrument. Three of them use logs to capture microprocesses. The chapters show that logs can be used both in open form for qualitative analyses and in standardized form for quantitative analyses.

The chapters demonstrate several advantages of logs. Logs have the potential to capture day-to-day behaviour in the context of school improvement, and it is precisely in that area that there is currently a lack of established instruments. Day-to-day behaviour (and other microprocesses) cannot be captured using most traditional questionnaires, because they were developed for cross-sectional designs. Moreover, qualitative studies seldom apply a methodology designed to carefully examine microprocesses longitudinally.

Logs have the advantage of having higher validity than traditional questionnaires that focus more on the measurement of abstracted activities from a longer period of time (Anusic, Lucas, & Donnellan, 2016; Ohly, Sonnentag, Niessen, & Zapf, 2010; Reis & Gable, 2000). Logs can provide better insights into day-to-day activities and their dynamics. This means that also shorter time periods and shorter intervals between the measurements can be examined. Both play an important role in investigation of the highly dynamic and very diverse school improvement processes frequently found in schools, such as initiation of changes, team building, the handling of pressing problems, and so on.

Exactly these processes must be investigated, if the aim is to better understand school improvement in the context of new governance and accountability measures. Data gathered with standardized logs can be analyzed using many established statistical methods for time series analysis (Hamaker, Kuiper, & Grasman, 2015; McArdle, 2009; Valsiner, Molenaar, Lyra, & Chaudhary, 2009). Furthermore, with sufficiently large samples and measurement points, logs allow multilevel analysis and thus the analysis of interaction effects between the different levels, such as between school, person, and time. One methodology that is particularly geared towards processes and dynamics of individuals, as presented by Oude Groote Beverborg et al. (Chap. 11), allows the analysis of regularity and stability of (the coupling between) microprocesses and improvement. Using qualitative logs that were sensitive to local and personal circumstances and Recurrence Quantification Analysis, they were able to analyze the extent to which differences in the regularity and frequency of teacher reflection in the context of workplace learning are connected with their own developments.

The more qualitative methodologies presented in this volume (Chaps. 9, 10, and 11) also allow to acquire more detailed findings on the extent to which attitudes,

orientations, and perspective towards school tasks and school improvement processes change. However, the particular challenge these kinds of studies face is the identification of substantial changes and to differentiate them from more random or insignificant developments. Therefore, the illustrative studies' log-based methodologies, as well as the corresponding conceptualizations and theories, need to be further developed and applied to different situations and school improvement contexts. This is particularly relevant in connection with questions pertaining to new governance and accountability measures. Previous research has insufficiently studied how teachers and school leaders, as well as other actors, react to external demands or monitoring outcomes, integrate them in their school practices (or not), and utilize them for teaching and student learning (or not). Commonly used questionnaires or interviews capture retrospective self-reports and are thus limited in tapping into ongoing improvement processes. In this regard, the methodological and theoretical developments presented in Chaps. 9, 10, 11, and 12 hold the promise of a substantial gain in knowledge and a significant broadening and deepening of understanding the connection between accountability and school improvement.

A prerequisite for the use of logs to capture behaviour in a day-to-day manner is the validity of the log itself. How logs can be validated ideally using observations and interviews is described in the contribution by Spillane and Zuberi (Chap. 9). Beyond that, there are additional challenges that must be tackled, because of the temporal nature of change and development in school practices, the role of actors' motivations or perspectives within school improvement processes, or monitoring procedures. A main keyword here is 'measurement invariance.' The contributions by Lomos (Chap. 4) and Sauerwein and Theis (Chap. 5) provide insight into analyses for testing measurement invariance using Multiple Group Confirmatory Factor Analysis (MGCFA). Although the analyses presented in these two contributions are based on cross-sectional data, MGCFA can be used to assess whether the meaning of a construct remains stable across different time points. In addition, MGCFA allows the examination of change in understanding of a construct itself or differences between groups in their (change of) understandings of a construct.

Especially regarding the interpretation of findings on measurement invariance (or measurement variance), however, there are a number of substantial research gaps. Measurement (in)variance can be technically determined, but the interpretation of such a finding depends on one's theory. A finding that points to measurement variance could – from a methodological viewpoint – indicate that longitudinal analysis should not be conducted. However, the finding could also indicate that the meaning of the items within a construct has changed over time for the participants. This is often the very goal of a school improvement measures, for instance, when the aim is to implement collegial cooperation or raise commitment. In the future, therefore, findings should be carefully considered on their methodological and theoretical merit, and separated using suitable methodologies when needed.

Also needed are measurement instruments that are specifically developed for empirically depicting the developmental courses of processes. This is particularly important for processes where development means not simply 'more of the same,' such as in the form of higher approval, intensity, and so on, but where the construct

itself changes. For example, with collegial cooperation, rudimentary cooperation is characterized simply by exchange of materials, whereas high-quality cooperation is characterized by co-constructive development of concepts and materials (Decuyper, Dochy, & Van den Bossche, 2010; Gräsel, Fußangel, & Pröbstel, 2006). Accordingly, forms of adaptive measurement could be developed in school improvement research, something that is being done for some time now in the area of competency assessment (Eggen, 2008; Meijer & Nering, 1999). Alternatively or concomitantly, researchers could work together with practitioners in common contexts to co-develop scales and the meaning of their intervals.

13.2 School Improvement as a Multilevel Phenomenon: The Meaning of Context for School Improvement

School improvement processes make up a complex phenomenon that takes place at different levels not only within the education system but also within schools. Accordingly, the notion of ‘context’ is quite complex.

As discussed in the contribution by Reynolds and Neeleman (Chap. 3), the improvement of schools and the underlying processes depend heavily on the social, socioeconomic, and cultural context of the school, as well as on the accountability modus that is implemented in the particular education system. In this sense, context refers to political, cultural, and social factors external to the school. Within schools, however, the organization (e.g. leadership) might be the context for teachers’ team learning, and consequently, teachers’ team learning can be understood as a context for teachers’ learning and teaching.

In the last 20 years, many empirical studies have shown that it is essential to consider these nested structures at the appropriate levels when investigating school improvement processes (see Hallinger & Heck, 1998; Heck & Thomas, 2009; Van den Noortgate, Opdenakker, & Onghena, 2005). However, there are several problems and challenges, particularly regarding the analysis of the multilevel structure of school improvement and the issue of how different contexts can be identified and taken into account. Several chapters in this volume discuss these points in detail.

First of all, the chapters in this volume that used logs in order to investigate day-to-day activities (for example, the contributions by Spillane and Zuberi and by Maag Merki et al.) point out that in school improvement research the hierarchical structure must be extended to include (at least) two further levels: daily activities and individual activities. The level of daily activities can then be considered as ‘nested in persons’, and the individual activities are then activities ‘nested in days’. With this, an extensive nesting structure of school improvement processes unfolds: individual activities, nested in days, nested in persons, nested in teams, nested in schools, nested in districts or regions, nested in countries. Development of the appropriate methodology and empirical assessment of this structure is challenging and future school improvement research could concentrate on that.

To take account of the hierarchical structure, hierarchical multilevel analyses have become the standard (e.g. Luyten & Sammons, 2010). Nevertheless, Schudel and Maag Merki (Chap. 12 in this volume) have critically discussed the existing practice of multilevel analysis. Although nested structures are taken into account in multilevel analysis, for instance through correction of standard errors, important information is lost with the common aggregation of data (which allows the use of information at higher levels). In addition, current research focuses solely on the group mean as a measure for shared properties. Variances in the aggregated properties or other parameters in the composition of these properties are thus overlooked. Therefore, as Schudel and Maag Merki mention, multilevel models in educational research have to consider the double character of groups: global group properties emerge from the group level and group composition properties emerge from the lower, individual level. Moreover, educational researchers have to take into account the possibility of both shared properties and configural properties of group compositions. In this way, the composition of the teaching staff, as well as the position of the individual within the teaching staff, can be regarded as an independent and process-relevant aspect of the multilevel structure, and the relation of either or both with individual teacher's actions and experiences can be examined. The use of the Group Actor-Partner Interdependence Model (GAPIM) allows a more differentiated modelling of, for instance, the frequently observed divergence in actors' perspectives on the implementation of reforms or their divergence in handling accountability requirements (e.g. interested and motivated teachers versus those who are opposed). Thus, the GAPIM allows a more valid investigation of how school improvement measures affect teachers' instructions and students' learning.

Further questions that could be interesting for both school improvement research and assessment of accountability processes are, for example: What dynamics emerge out of which (properties) of group compositions? What changes in composition are affected by school improvement measures (such as measures to develop a shared educational understanding, to reach an agreement on guiding principles, and so on)? Can different developmental courses in schools be explained by group composition properties? What aspects of the composition of the teaching staff are important for the success of school improvement measures?

Ng (Chap. 7) argued for another approach to identifying school-internal context conditions: social network analysis. This methodology has only been adopted in a few studies up to now (Moolenaar, Slegers, & Daly, 2012; Spillane, Hopkins, & Sweet, 2015; Spillane, Shirrell, & Adhikari, 2018). Social network analysis allows examination of the social structure of school teams and investigation of how this structure affects teachers' practices and the school's improvement processes. A clear gain over other methodologies is that the loosely coupled structures of schools (Weick, 1976) can be made visible. As such, formal and informal team structures, as well as densities of ties within teams and with other actors, can be investigated with respect to sustainable school improvement. In addition, the methodology also makes it possible to compare individual schools, which may uncover explanations for school-specific developmental trajectories of students.

Vanblaere and Devos (Chap. 10) investigated the effect of context from yet another perspective. Their focus was on a school-specific innovation, which they assessed with qualitative teacher logs over the course of a year in four primary schools, which were characterized as either a high or a low professional learning community (PLC). With such qualitative logs, it is possible to assess developments in each separate school, while taking different starting conditions (low and high PLC) into account. When using such unstandardized logs, developmental courses and events can be captured that had not been anticipated in advance.

The presented studies open up new perspectives to include context in the study of school improvement and school practices. However, many aspects are still not taken into sufficient consideration. In particular, investigations of how aspects of contexts affect actors should be extended with detailed assessments of the extent to which actors themselves change their contexts through their perceptions of, and actions in, those contexts (Giddens, 1984). This continuous interaction would require a longitudinal design and methodology in addition to multilevel methodology, and this has not been considered enough in previous research. Measurement instruments must therefore be sufficiently sensitive regarding differences in contexts but also regarding the identification of changes (at different levels), which is a double challenge. Beyond that, more differentiated investigation is needed on the extent to which school improvement strategies are dependent on certain contexts to be functional for sustainable development, or on what strategies are particularly productive for schools with either high or low school improvement capacities. This raises the issue of generic or specific school improvement processes and success factors (Kyriakides, 2007).

13.3 Indirect and Reciprocal Effects

School improvement is a complex process in which many processes (e.g. leadership actions, decisions and actions of several teams, and individual teachers) are involved over time. This process takes place at different levels (school level, team level, classroom level). From this point of view, school improvement processes usually have direct and indirect effects. Twenty years ago, Hallinger and Heck (1998) already pointed out for school leadership research that ignoring indirect effects impacts the validity of findings on the effect of school principals' actions on student achievement. The same can be assumed also for school improvement processes and for processes connected with accountability requirements and reforms. Due to the number of factors involved in those processes and the resulting number of hypothetically possible direct and indirect effects, it is not possible to assess all direct and indirect effects simultaneously (for example using structural equation models). Here it is important to carefully consider what direct and indirect effects should be included in the theoretical and the empirical model, and, where needed, to test individual paths one after the other and in advance.

Indirect relations were addressed in the contribution by Ng (Chap. 7). Ng describes an example of a social network analysis that was used to identify heterarchical paths of decision-making processes in schools, even though the structure of the school was organized hierarchically. Social network analyses are suited to identify for individual schools via which and via how many others persons are connected in a network. These relationship structures represent the potential to spread content. In this regard, communication and decision paths as well as cooperation and power structures, for example, can be analysed as microprocesses with social network analysis. In addition to indirect effects, social network analysis can also be used to identify reciprocal effects, and in which schools teachers are connected only unidirectionally (person A chooses person B, but person B does not choose person A) or mutually and thus reciprocally (person A chooses person B, and person B chooses person A).

Indirect relations were also identified by Maag Merki et al. (Chap. 12). Multilevel analysis of the log data revealed that the relation between teachers' ratings of the day's activities and their daily satisfaction varied school-specifically and that it was moderated by teachers' interests in assessment and further development of their own teaching practices. Although these findings need to be tested in larger samples, they show the potential of log data to reveal differential and indirect effects. Complementary qualitative analyses could provide greater depth, such as was done in the study by Vanblaere and Devos (Chap. 10). In this way, explanations can be found that help to further develop theoretical models.

13.4 Variety of Meaningful Factors

To understand and assess school improvement processes, it is important to take a broad view of possible dimensions, structures, processes, and effects. Nevertheless, current school improvement research has strongly built on well-established dimensions and empirical findings (such as leadership practices or cooperation), which resulted in limited variability in research focus, and this has possibly limited development of more fully understanding the mechanisms involved in school improvement. An interesting extension of research on school leadership is presented in the contribution by Lowenhaupt (Chap. 8). In the study, the focus is on a linguistics method for analysing the rhetoric of school leaders. Lowenhaupt discovered that the rhetoric that school leaders use varies, and that rational, ethical or affective aspects are emphasized depending on the situation. As such, school leaders aim to initiate or influence development processes and school practices by differentiating their rhetoric. It would be interesting to investigate how differing rhetorical means affect teachers' motivation or interest in reflecting on their own practice in terms of quality development, how rhetorical means covary with individual characteristics, or how their availability and use change over time. The methodology can be linked to neo-institutional theories (DiMaggio & Powell, 1983/1991) or micropolitical theories for assessment of organizations (Altrichter & Moosbrugger, 2015). As such, it

allows differentiated analysis of power structures, negotiation processes on goals, values, and norms, and it can provide a better understanding of why school reforms do not, or only partially, achieve desired aims. In this sense the methodology presented holds potential for future school improvement research and for studies assessing intended and unintended effects of accountability approaches.

13.5 Concluding Remarks

The illustrative studies in this volume show how innovative methodologies can enrich school improvement research and help further development thereof. Taken together, they also provide an overview that can be used to systematically select the kind of methodology that fits a certain aspect of school improvement best. Moreover, we think that multimethod designs in which the presented methodologies are combined with other, especially qualitative, methodologies are very promising to better understand the complex interplay between actors' subjective meanings, their attributions, motivations, and orientations (e.g. Weick, 1995), individual and collective actions, and school structures and educational systems.

The methodologies presented in this volume for studying school improvement processes in the context of complex education systems cannot claim to revolutionize school improvement research, especially because the contributions could only selectively address previous research gaps. In addition, investigation of, for instance, differential paths and nonlinear trajectories could not be included. Still, we hope that with the presented innovative methodologies and designs, as well as the resulting new perspectives, we have provided inspiration for the study of school improvement as a multilevel, complex, and dynamic phenomenon. Future studies on key aspects thereof will provide a deeper understanding of school improvement in the context of societal and professional demands, and this will have a positive effect on the quality of school organisation, instruction, and ultimately on student learning.

References

- Altrichter, H., & Moosbrugger, R. (2015). Micropolitics of schools. In J. D. Wright (Ed.), *International encyclopedia of the social & behavioral sciences* (Vol. 21, 2nd ed., pp. 134–140). Oxford, UK: Elsevier.
- Anusic, I., Lucas, R. E., & Donnellan, M. B. (2016). The validity of the day reconstruction method in the German socio-economic panel study. *Social Indicators Research*, 1–20. <https://doi.org/10.1007/s11205-015-1172-6>
- Decuyper, S., Dochy, F., & Van den Bossche, P. (2010). Grasping the dynamic complexity of team learning: An integrative model for effective team learning in organisations. *Educational Research Review*, 5(2), 111–133.
- DiMaggio, P. J., & Powell, W. W. (1983/1991). The Iron cage revisited: Institutional isomorphism and collective rationality. In W. W. Powell & P. J. DiMaggio (Eds.), *The new institutionalism in organizational analysis* (pp. 63–82). Chicago, IL: University Chicago Press.

- Eggen, T. J. H. M. (2008). Adaptive testing and item banking. In J. Hartig, E. Klieme, & D. Leutner (Eds.), *Assessment of competencies in educational contexts* (pp. 215–234). Göttingen, Germany: Hogrefe.
- Feldhoff, T., Radisch, F., & Klieme, E. (2014). Methods in longitudinal school improvement research: State of the art. *Journal of Educational Administration*, 52(5).
- Giddens, A. (1984). *The constitution of society: Outline of the theory of structuration*. Oakland, CA: University of California Press.
- Gräsel, C., Fußangel, K., & Pröbstel, C. (2006). Lehrkräfte zur Kooperation anregen – eine Aufgabe für Sisyphos? *Zeitschrift für Pädagogik*, 52(6), 205–219.
- Hallinger, P., & Heck, R. H. (1998). Exploring the principals' contribution to school effectiveness: 1980–1995. *School Effectiveness and School Improvement*, 9(2), 157–191.
- Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, 20(1), 102–116.
- Heck, R. H., & Thomas, S. L. (2009). *An introduction to multilevel modeling techniques* (Quantitative methodology series, 2nd ed.). New York, NY: Routledge.
- Kyriakides, L. (2007). Generic and differentiated models of educational effectiveness. In T. Townsend (Ed.), *International handbook on school effectiveness and improvement* (pp. 41–56). Dordrecht, The Netherlands: Springer.
- Luyten, H., & Sammons, P. (2010). Multilevel modelling. In B. P. M. Creemers, L. Kyriakides, & P. Sammons (Eds.), *Methodological advances in educational effectiveness research* (pp. 246–276). Abingdon, UK/New York, NY: Routledge.
- McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology*, 60, 577–605.
- Meijer, R. R., & Nering, M. L. (1999). Computerized adaptive testing: Overview and introduction. *Applied Psychological Measurement*, 23(3), 187–194.
- Moolenaar, N. M., Slegers, P. J. C., & Daly, A. J. (2012). Teaming up: Linking collaboration networks, collective efficacy, and student achievement. *Teaching and Teacher Education*, 28(2), 251–262.
- Ohly, S., Sonnentag, S., Niessen, C., & Zapf, D. (2010). Diary studies in organizational research. An introduction and some practical recommendations. *Journal of Personnel Psychology*, 9(2), 79–93. <https://doi.org/10.1027/1866-5888/a000009>
- Reis, H. T., & Gable, S. L. (2000). Event-sampling and other methods for studying everyday experience. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 190–222). New York, NY: Cambridge University Press.
- Spillane, J. P., Hopkins, M., & Sweet, T. M. (2015). Intra- and interschool interactions about instruction: Exploring the conditions for social capital development. *American Journal of Education*, 122(1), 71–110.
- Spillane, J. P., Shirrell, M., & Adhikari, S. (2018). Constructing “experts” among peers: Educational infrastructure, test data, and teachers' interactions about teaching. *Educational Evaluation and Policy Analysis*, 40(4), 586–612.
- Stoll, L., & Fink, D. (1996). *Changing our schools: Linking school effectiveness and school improvement*. Buckingham, UK: Open University Press.
- Valsiner, J., Molenaar, P. C., Lyra, M. C., & Chaudhary, N. (Eds.). (2009). *Dynamic process methodology in the social and developmental sciences*. New York, NY: Springer.
- Van den Noortgate, W., Opdenakker, M. C., & Onghena, P. (2005). The effects of ignoring a level in multilevel analysis. *School Effectiveness and School Improvement*, 16(3), 281–303.
- Weick, K. E. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, 21, 1–19.
- Weick, K. E. (1995). *Sensemaking in organizations*. London, UK: Sage.

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

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.