Thillmann, Katja; Bach, Anabell; Wurster, Sebastian; Thiel, Felicitas

School-based staff development in two federal states in Germany

School-based staff development in two federal states in Germany

Katja Thillmann and **Anabel Bach***: Schulpädagogik/Schulentwicklungsforschung, Freie Universität Berlin, Berlin, Germany
Sebastian Wurster: Institut zur Qualitätsentwicklung im Bildungswesen (IQB), Humboldt-Universität zu Berlin, Berlin, Germany,
Felicitas Thiel: Schulpädagogik/Schulentwicklungsforschung, Freie Universität Berlin, Berlin, Germany
*Corresponding author: Anabel Bach (e-mail: anabel.bach@fu-berlin.de)

Abstract

Purpose – In Germany up until now, there has been very little research on staff development in schools. The purpose of this paper is to comprehensively assess school-based staff development and to describe the interplay between different instruments of staff development (e.g. classroom observations, development discussions) at the school level.

Design/methodology/approach – Considering that different constellations of organizational management tools may be differentially effective in different contexts (see Mintzberg, 1983/1992), an approach that takes a combination of different staff development instruments into account was chosen. Data were gathered from principals of primary and secondary schools in two federal states of Germany. Using regression, cluster analysis, and analysis of variance, the authors examined different instruments and patterns of staff development used in everyday school practice and determined how these affected the professional development of teachers.

Findings – Five staff development patterns could be identified. With regard to the extent of professional development activities of teachers, these patterns have been proven to have a different impact. Furthermore, the use of the different staff development patterns seems to be heavily dependent on the type of school.

Research limitations/implications – Further research would be needed that examines if the three most relevant staff development patterns identified in this study can also be proven to be effective with regard to somewhat "harder" criteria than the extent of professional development activities of teachers. Such criteria could be teachers' teaching skills or even student achievement.

Originality/value – The current study is the first to examine staff development in German schools systematically. The results provide some good leads for further studies in this area.

Keywords: Staff development, School autonomy, Professional development, Organizational management, Classroom observation

1. Introduction

In Germany, staff development has gained a high significance since the introduction of enhanced school autonomy. Apart from greater opportunities for staff recruitment, the corresponding legal changes in education also imply structural changes with regard to the relationship between principals and their staff. Legally, principals are managers, thus creating a vertical hierarchy in schools, whereby principals have the managerial authority toward their teaching staff. One key area of principals' managerial duty is staff development (see Huber, 2013).

Principals can make use of a variety of instruments and methods (e.g. classroom observations, development discussions, professional development planning) for systematic staff development. Simultaneously with the introduction of the new approach to school governance in Germany, the importance of approaches to collaborative professional development (e.g. student surveys, peer classroom observations) for teachers was emphasized. If systematically implemented, such approaches can also be part of the school-based staff development.

If and in what way different instruments and procedures of staff development are actually used in practice, and to what degree they differ in their impact, has so far not been examined systematically in Germany. The current study addresses this issue. In this study, we regard the professional development activities of teachers as a proxy indicator for the impact of staff development.

First, we introduce the concept of staff development and its relevance in a school context in Germany (2). Second, we describe different instruments of staff development in German schools (2.1). Then, we briefly review the relevant research literature regarding school-based staff development (2.2). Here, we give a short overview of different research traditions in this area, followed by a summary of previous research findings most pertinent to our study. Subsequently, we present our research questions (3) and methods (4), our results (5), and the discussion and future directions (6).

2. Theoretical framework

In Germany, staff development is a relatively new task for schools, and its implementation so far has been rarely researched systematically. This is mostly due to the great heterogeneity and fuzziness in the concept of staff development itself (see Becker, 2009, p. 3). The literature features a variety of different definitions and taxonomies that often reflect quite diverse perspectives (see Becker, 2009, p. 5). There is some consensus that "staff development involves planned measures to increase one's professional competence" (Holling and Liepmann, 2007, p. 346, translated by the authors). Such measures can be implemented at different levels: they can be targeted at the development of individual competencies or at the development of the organization as a whole (Holling and Liepmann, 2007, p. 347). Professional development measures for teachers are at the core of school-based staff development (see Meetz, 2007, p. 94). However, coordination is necessary for a systematic staff development of teachers that takes individual as well as organizational development needs into account: "Harsh lessons from the past have taught educators that fragmented, piecemeal approaches to professional development do not work" (Guskey, 2000, p. 19). A coordinated approach can be realized in different ways: for instance by principals identifying and addressing the development needs of their teachers (e.g. through classroom observations, development discussions, and individual performance reviews), by teachers determining their development needs in a collaborative professional manner (e.g. through self-evaluation, peer classroom observation), or through school-wide staff development programs that are based on a systematic needs assessment (see Thillmann, 2012, for an overview of coordination in school organizations).

2.1 Instruments of staff development in schools

In the following, we give a brief overview of the main instruments of staff development in schools, which were implemented in the German federal states as part of the increased autonomy of schools

regarding their staff. Using expert ratings of school law, the Action Committee on Education (Vereinigung der Bayerischen Wirtschaft e.V., 2010) attested that schools in most states have extensive autonomy with regard to their staff development. Legally, principals are to fulfill new tasks like classroom observations, performance reviews, development discussions, as well as setting up school-wide staff development programs.

For purposes of staff development in schools, particular emphasis is placed on classroom observations and performance reviews. Staff development needs to be carried out on a basis of comprehensive information regarding the current competencies and skills and/or specific development needs of staff (see Meetz, 2007, p. 23). Classroom observations play an important role for performance appraisals. Legally, such classroom observations are to be conducted by principals once every four to five years as part of the performance appraisal process. Other instruments involve the teachers as active participants in the review process (see Buhren and Rolff, 2009, p. 36). Accordingly, the idea of collaborative professional development has been promoted by education policy. For instance, peer classroom observations have been defined as an indicator of school quality (see e.g. Senatsverwaltung für Bildung, Wissenschaft und Forschung, 2013) and teacher self-evaluation procedures have been made compulsory[1]. Collaborative professional development activities can be counted as part of systematic school-wide staff development strategies if schools have systematically incorporated them (e.g. into their school improvement plan), and they are thus applied regularly.

Development discussions are another newly implemented instrument of systematic staff development in some of the German Federal States. Development discussions are not be confused with appraisal interviews. According to Buhren and Rolff (2009, p. 64), development discussions exclusively serve a development purpose rather than an appraisal purpose. Indeed, development discussions can be used to feedback the results of classroom observations or performance reviews, but they explicitly involve the teachers' perspective. Individual performance reviews and development discussions are regarded as a low stakes measure. Performance goals are to be set in a joint manner, whereby the focus is clearly on the development potential of teachers (for an overview of appraisal for the purpose of accountability vs development see Craft, 2000, pp. 35-36).

A further instrument for a systematic implementation of staff development is a professional development program. Such programs fix goals and measures for staff development in schools for a longer period of time (see Kieser and Kubicek, 1983, p. 112). In contrast to individual development discussions professional development programs define measures for the whole school: they are often linked to specific school development initiatives, such as improving reading skills of poor readers or implementing standards-based teaching (see Burkard, 2000). They systematically describe a set of measures to extend the competencies of the whole staff or particularly of the part of the staff which deals with new tasks or challenges.

2.2 Research findings

Due to the heterogeneous nature of this area of research, we first focus on general findings concerning staff development from a national and international perspective (2.2.1) and then report specific findings that are of particular relevance to the current study (2.2.2). In the second part, we focus on reviewing research about the prevalence of the principal instruments of staff development we outlined in Section 2.1 and further limit our presentation to findings from Germany, since here the new measures and approaches to school governance have been implemented under specific circumstances. For example, as described above, in Germany the instrument of "classroom observations" is usually only applied in the context of formal performance appraisals. In comparison, teacher appraisals in the USA, for example, have been implemented as part of the accountability movement and are thus conducted under completely different conditions (see Ovando and Ramirez, 2007).

Furthermore, there is a distinct difference between the tasks and responsibilities of German principals when compared internationally. The enhanced autonomy of German principals, for example with

regard to staff selection or spending (e.g. for particular staff development activities), is relatively low when compared internationally (Rosenbusch et al., 2006, p. 9). The responsibilities of German principals are mostly characterized by pedagogical activities (high proportion of classes taught). In the USA by contrast, management and leadership activities are predominant tasks (Wissinger, 2002, p. 55). Further, in comparison to the UK or USA, for example, Germany is also lacking a systematic training of principals. Rosenbusch and Huber (2001) for instance showed that while all German states offer further education trainings, these are generally very dissimilar and "not well thought-out" (p. 10) with regard to their contents and conceptualization. Because of the specific circumstances in Germany, an extensive international comparison regarding the application of different instruments of staff development seems of limited value to us. Moreover, due to cultural-specific effects (see Ngo et al., 1998), Sonntag and Stegmaier (2006, p. 406) advice extreme caution when transferring research results regarding the application and effects of staff development measures to other cultures.

2.2.1 Research on staff development in schools

Meetz (2007, pp. 137-154) illustrates the range of research in this area using four exemplary studies that examined the practice of teacher professional development in Germany (e.g. participation in professional development activities, content of professional development activities), career paths and performance incentives (e.g. salary bands, promotion opportunities), as well as the prevalence of staff development measures like performance appraisals or development discussions. Appius et al. (2012, p. 124)[2] conceptualize staff development as a core task of principals and correspondingly point out studies that focus on the managerial role of principals. They researched the professional ethos of principals, their subjectively estimated scope on decision making regarding staff development, and the teachers' perspective on the implementation of staff development measures (Appius et al., 2012, p. 125). In sum, German-speaking research results have so far been of rather limited value (Terhart, 2010, p. 258) and do not allow for inferences regarding potential effects on teachers' classroom practices or student achievement.

In contrast, research at the international level, as described below, has produced a range of more in-depth results for some aspects of staff development. However even internationally this area is still lacking some theoretical and empirical insights (Steger Vogt, 2013, pp. 41-48). There are for instance few empirical findings from Anglo-American studies regarding the effectiveness of staff development measures (Dubs, 2005; Desimone, 2009). Some studies examined the effectiveness of formal as well as informal (on-the-job) learning arrangements (e.g. peer classroom observation, peer coaching), whereby quite different aspects of effectiveness were often studied (Lipowsky, 2010). There are relatively many studies using participating teachers' ratings to determine factors that predict satisfaction with and acceptance of professional trainings (Smith and Gillespie, 2007). If, however, positive evaluations of professional trainings by participants can predict a change in their classroom practices, remains to be proven. First results, which are in line with findings from research on trainings generally, do not point in that direction (Goldschmidt and Phelps, 2010). Other studies examined changes in teachers' cognitions (e.g. Franke et al., 2001), classroom practices (e.g. Desimone et al., 2002), and associations with student achievement (e.g. Parr et al., 2007).

Another research strand, which so far has been very weakly represented in the German studies, addresses the importance of performance reviews for the professionalization of teachers. Such studies are mostly situated in the context of leadership research (see Leithwood et al., 2004) and are aimed at determining specific traits of principals that predict successful appraisal and feedback processes (e.g. Ovando and Ramirez, 2007; Tuytens and Devos, 2011). The international research on principals has long-since highlighted the importance of classroom evaluations conducted by principals under the heading of "instructional leadership" (e.g. Goldman, 1998; Spillane et al., 2004). Furthermore, studies that explore "management by wandering around" (see Brooks et al., 2007) or "classroom walkthrough" (see David, 2008; Kachur et al., 2010), whereby principals visit classrooms (often

unannounced), are of particular significance, especially with respect to the staff development instrument "classroom observations" (see Blase and Blase, 1998, pp. 105-111). In their meta-analysis, Robinson et al. (2008) conclude that with regard to student achievement instructional leadership is more effective than transformational leadership. Whereby, among other factors, regular classroom evaluations are closely related to the improvement of student achievement (Robinson et al., 2008). This finding was also supported by a meta-analysis about the relationship between leadership and student achievement, which produced similar results (Waters et al., 2003).

Internationally as well as nationally, studies are lacking which empirically assess staff development in schools as a holistic concept and examine the interplay between different measures at the school level. Overall, research on staff development in schools has mostly been limited by a narrow focus on the effectiveness of separate measures (e.g. classroom observations, see Downey et al., 2010, for the international context), or single aspects (e.g. professional development of teachers in learning communities, see Gräsel et al., 2006, for the national, Stoll and Louis, 2007, for the international context; or feedback processes, see Smither et al., 2005).

2.2.2   Research on different instruments of staff development in German schools

Findings regarding classroom observations and development discussions generally show that at most schools classroom observations and development discussions are not yet being used as a common practice. Bonsen et al. (2002, pp. 74-75), for example, showed in an explorative study examining 30 schools that in the majority of those schools, classroom observations were only conducted if there was a specific reason (e.g. upcoming promotion). Similarly, results from the PISA 2003 additional survey of school principals also demonstrated that only a quarter of the respondents conducted regular annual classroom observations (Senkbeil et al., 2004, p. 310).

Results from an evaluation of a Berlin pilot project regarding school improvement paint a similar picture. According to this study, only one in five teachers reported having their classes observed by their principal in the preceding two years. Similarly, Avenarius et al. (2006, pp. 118-119) report that only between 15 and 32 percent of teachers have experienced structured development discussions conducted by their principals. An explorative study involving ten secondary schools in North Rhine-Westphalia (Meetz, 2007, p. 182) also documents that only half of the surveyed teachers have had development discussions conducted by their principals in the preceding two years. Even a more recent study by Appius et al. (2012) found that in the federal state of Baden-Wuerttemberg and the Swiss cantons of St. Gallen and Lucerne, both development discussions as well as performance agreements were conducted irregularly in most schools. Results from an evaluation of a pilot project in North Rhine-Westphalia[3] paint a more positive picture. According to the surveyed principals, 79 percent of schools conducted development discussions and 63 percent conducted performance agreements. Classroom observations, on the other hand, were markedly less common (Klemm and Meetz, 2008, p. 176).

Regarding the practice of staff development programs in schools, Appius et al. (2012, p. 132) note that staff development activities were only partly implemented within the framework of such programs. Likewise, the results from the evaluation of the Berlin pilot project on school improvement showed that at the time of the evaluation only half of the schools had a staff development program in place (Avenarius et al., 2006, p. 122).

Such results are also reflected in the findings of school improvement plan analyses. The content analyses of Hamburg school improvement plans conducted by Holtappels and Müller (2002, p. 222) found only 37 percent of schools to feature a systematic staff development program in their improvement plans. As part of the Progress in International Reading Literacy Study (PIRLS), Mohr (2006, p. 91) found that only 10 percent of the school improvement plans explicitly mentioned developing teacher competencies as a main area of school improvement.

Again, results from the evaluation of the pilot project in North Rhine-Westphalia paint a much more positive picture. In 2007, 90 percent of the evaluated principals stated that they implemented a staff development program. Further, a needs assessment for professional development was conducted by the majority of the schools (64 percent) regularly and systematically (77 percent) (Klemm and Meetz, 2008, p. 176). However, generally it is questionable to what degree such findings from pilot project schools can be generalized to regular schools.

Similar results can be found with respect to the practice of collaborative professional development among teachers. In fact, according to the principals surveyed by Bonsen et al. (2002, p. 84), more than half reported forms of classroom feedback (e.g. by student surveys or peer classroom observations), yet it remains unclear to what degree these instruments are applied systematically or regularly. Appius et al. (2012, p. 130) also found that peer classroom observations were conducted irregularly. Furthermore, using self-evaluation instruments does not seem to be an established practice yet. According to the PISA additional survey of school principals, only one fifth of schools review their work regularly using self-evaluation (Senkbeil et al., 2004, p. 307).

However, data from the evaluation of the pilot project in North Rhine-Westphalia show a significant increase between 2003 and 2007 in the frequency of self-evaluations (e.g. regular student feedback from 9 to 37 percent) (Feldhoff et al., 2008, p. 185). Yet again, the applicability of such results from pilot project schools to regular schools has to be called in question.

Taken together, the findings suggest that staff development has so far not been practiced systematically but is rather carried out in single measures (Appius et al., 2012, p. 124; Meetz, 2007, p. 12). However, the reviewed results do not give any indication about the interplay between the different instruments at the school level or the effectiveness of different staff development instruments.

Considering that different constellations of organizational management tools may be differentially effective in different contexts (see Mintzberg, 1983/1992), an approach that takes a combination of different staff development instruments into account seems warranted and advantageous.

3. Research questions and methods

Our research aims at addressing the above mentioned research gaps. Specifically, we address the following questions:

RQ1. To what degree are different instruments of staff development (classroom observations, development discussions, staff development programs, and forms of collaborative professional development) applied in German schools?

RQ2. To what extent do teachers participate in professional development activities?

RQ3. To what degree do staff development instruments affect the participation in professional development activities of teachers?

RQ4. To what extent are different instruments used in practice associated? Are there typical combinations of instruments, and can thus different patterns of staff development be identified?

RQ5. Are different patterns of staff development associated with context factors at the school level (e.g. characteristics of principals, type of school)?

RQ6. Which staff development pattern is of significant importance for the professional development activities of teachers?

4. Method

4.1 Sample

This research addressed all principals of primary and secondary schools in the federal states of Berlin and Brandenburg in Germany. The survey was carried out online between November 2011 and

February 2012 as a first part of a wider study funded by the Federal Ministry of Education and Research. In total n=327 principals participated in this survey. The response rate was 27 percent. The sample in both states is representative with regard to the type of school and school region.

4.2 Measures

Instruments of staff development in schools (see Table I): the instruments of staff development described above were assessed as follows. Regular classroom observations was assessed with one item and a dichotomous yes/no response format. Development discussions and staff development programs were assessed using scales with four items. The discrimination power of the development discussions scale ranged from r_{it}=0.40 to 0.45 and was sufficient. Furthermore the discrimination power of the staff development programs scale ranges from r_{it}=0.37 to 0.44 and was sufficient as well. The variable collaborative professional development comprises a mean index of four items (peer classroom observation, parallel student tests, oral student feedback rounds, paper-pencil student surveys regarding the quality of classes).

Table I: Descriptive statistics of the independent variables

Variable	No. of Items	Example Item	α	M	SD	N
Regular classroom observations [a] Single item	1	---	---	.51	.50	245
Development discussions [b] Scale	4	For what purposes do you conduct development discussions? - I mainly use development discussions to agree on measureable performance targets with my teachers.	.631	4.26	.89	243
Staff development programs [b] Scale	4	How would you describe the staff development program at your school? - We have aligned our staff development program purposefully with our educational focus.	.609	4.37	.88	245
Collaborative professional development [c] Mean index	4	Please rate how many of your teaching staff regularly conduct the following types of classroom evaluation? - Peer classroom observation	---	3.13	.85	312

[a] dichotomous response format (0-no. 1-yes)
[b] 6-point scale ranging from 1 = "not true at all" to 6 = "exactly true"
[c] 6-point scale ranging from 1 = "nobody" to 6 = "everybody"

Participation in professional development activities: Teachers' participation in professional development activities was assessed by having principals rate what percentage of their teachers participated on average in none, one, two, or more professional development activities per school year. Context factors (see Table II). The following context factors were included in the analysis: federal state, type of school, size of school (number of teachers), age of the principal, average age of the

teachers, and the proportion of immigrant students. The age of the teachers, as well as the proportion of immigrant students and the size of the school were assessed using an open-ended response format. The age of the principal was assessed using a five-point response format. The variables federal state (0=Brandenburg, 1=Berlin) and type of school (0=primary school, 1=secondary school) were scored dichotomously. In-depth analyses, whereby secondary schools were further distinguished into Gymnasium[4] and other secondary schools, were also conducted.

Table II: Descriptive statistics of the context factors

Variable	Item	M	SD	%	N
Size of school Single item	How many teachers are currently employed at your school?	30.72	22.32	-	326
Average age of the teachers Single item	What is the average age of the teachers at your school?	49.49	3.57	-	323
Proportion of immigrant students Single item	What proportion of immigrant student does your school feature?	13.51	23.27	-	326
Federal state Single item	Brandenburg Berlin	- 	- 	66.06 33.94	327
Type of school Single item	Primary School Secondary School	- 	- 	62.39 37.61	327
Age of the principal Single item	30 yrs. or younger 31 to 40 yrs. 41 to 50 yrs. 51 to 60 yrs. 60 yrs.or older	- 	- 	0 1.23 28.70 57.72 12.35	324

4.3 Analysis

The analysis was undertaken in four steps. First, descriptive analysis was applied to examine to what degree different instruments of staff development are used in schools, and to what extent teachers participate in professional development activities. Second, a multiple regression analysis was conducted to determine the extent to which professional development activities of teachers could be predicted by the use of different staff development instruments. In line with our research question, we included all variables related to professional development (regular classroom observations, development discussions, staff development programs, and collaborative professional development) in our model as well as the control variables federal state and type of school. We used multiple imputation to deal with missing data in the regression analysis, applying predictive mean matching (Little, 1988), and preparing 50 imputed data sets. The range of the amount of missing data varied from 0 (federal state) to 33 percent (regular classroom observations). The examination of the missing patterns show no hints that the missing at random assumption is violated, thus multiple imputation is suitable.

Following the assumption that effective staff development depends (in part) on a specific combination of different instruments (see Mintzberg, 1983/1992), in a third step, cluster analyses were conducted to identify typical staff development patterns used by schools. The cluster analysis was based on the use of the four staff development instruments[5]. First, a hierarchical cluster analysis according to the Ward-method was conducted. To determine the appropriate number of clusters, we took into account the heterogeneity measure, the internal homogeneity (F-values), as well as the stability of the different solutions (cross-validation). We also used the interpretability of the solution as a criterion. The solution was then optimized using the k-mean procedure (Bortz, 2005, p. 575)[6].

In a fourth step, we conducted analyses of variance (for the metric variables) and χ^2 tests (for the nominal variables) in order to determine differences between the different staff development patterns with regard to different context criteria (e.g. type of school, size of school) and the professional development activities of teachers. η^2 served as an effect size for the strength of the influence of the cluster assignment on the metric dependent variable[7]. To evaluate pairwise group differences, we also calculated Cohen's d[8]. For a global evaluation of group difference regarding the nominal variables, the effect size ω^2 was used (see Rasch et al., 2010, p. 181)[9].

5. Results

(1) Addressing our first research question about the degree to which different instruments are applied in German schools, we found that about half of the surveyed principals (51 percent) conducted regular classroom observations (at least once every one or two years) outside of the formal appraisal process. Nearly two-thirds of the surveyed principals (65 percent) indicated that they conducted regular development discussions with their teachers. A majority of principals used these development discussions to provoke teachers' self-reflection (M=4.68, SD=1.16)[10] or to agree on performance targets (M=4.07, SD=1.40)[11]. Also, according to the surveyed principals, systematic professional development planning took place in many schools: the majority tended to agree that they purposefully informed themselves about their teachers' classes in order to determine the teachers' individual development needs (M=3.98, SD=1.35), or that they deliberately aligned staff development plans with the school program (M=4.20, SD=1.40). Regarding different forms of collaborative professionalization, however, so far these have rarely been implemented systematically in practice. So, according to the surveyed principals, parallel tests and student feedback rounds were regularly conducted by the majority of teachers (over 80 percent) in only 11 percent of the schools. Paper-pencil student surveys and peer classroom observations were regularly conducted by teachers at only 4 percent of the surveyed schools.

(2) Concerning our second research question about the extent to which teachers participate in professional development activities, our results show that principals estimated that on average about 50 percent of teachers at their schools participated in more than two professional development activities per year. On the other hand, 9 percent of teachers did not participate in any professional development activities. On average, the extent of participation in professional development activities was significantly higher in schools in Brandenburg when compared to Berlin schools (see Figure 1).
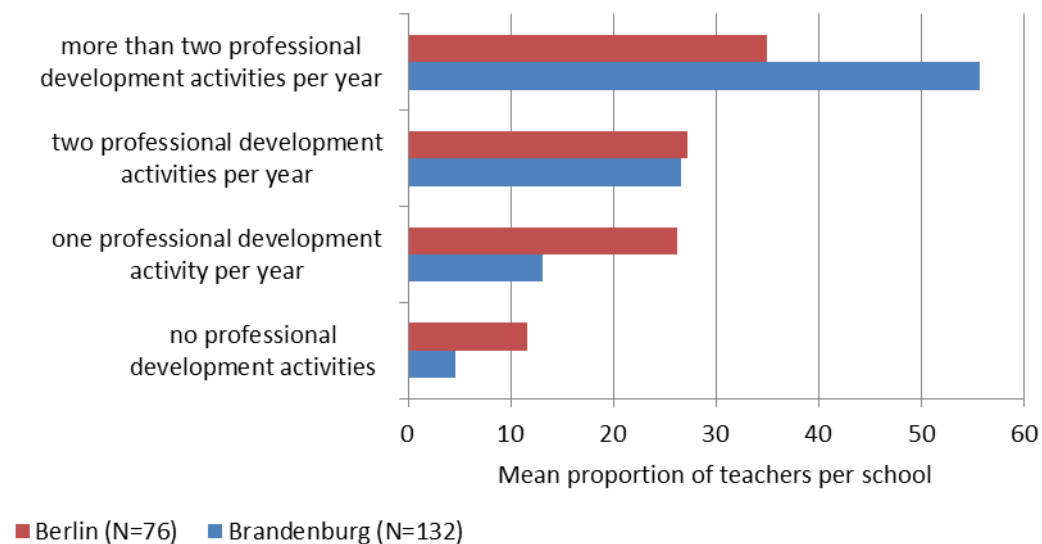
Figure 1. Professional development activities in Berlin and Brandenburg

(3) Our third research question was about the significance of staff development instruments for teachers' participation in professional development activities. Primary school teachers participated in significantly more professional development activities than secondary school teachers. While especially the difference in states (β=−0.18*) and the type of school (β=0.21***) seemed to influence the extent of professional development activities, the use of staff development measures also played a significant role. Having systematic professional development programs in place (β=0.18*) and conducting regular classroom observations (β=0.18*) seemed to be particularly important. In total, 33.1 percent of variance in the extent of participation in professional development activities could be explained by the independent variables (see Table III).

Table III: Results of the multiple regression analyses predicting the extent of participation in professional development activities (N=327)

	β	SE	95% CIs
Regular classroom observations	.18*	.07	[.04, .31]
Development discussions	.03	.08	[-..14, .18]
Staff development programs	.18*	.09	[.00, .37]
Collaborative professional development	.13	.08	[-.06, .31]
Federal state	-.18*	.07	[-.34, -.03]
Type of school	.21***	.06	[.10, .32]

Notes. R^2=.33; *p<.05, **p<.01, ***p<.001.

(4) Our fourth question regarding the different instruments and patterns of staff development, we examined using the Ward's method. First, a five-cluster solution was determined. The heterogeneity measure, however did not indicate a definite solution. As a further criterion, the analysis of F-values for different cluster solutions (2-6-cluster solutions) was used, according to which the 2-5-cluster solutions were homogenous (all F-values<1). A comparison of the cross-validations of all four solutions with the respective optimized k-means solutions showed the best match and thus the highest stability for the three- and five-cluster solutions (κ=0.84, 0.83, respectively). The five-cluster solution could be interpreted best (see Figure 2).
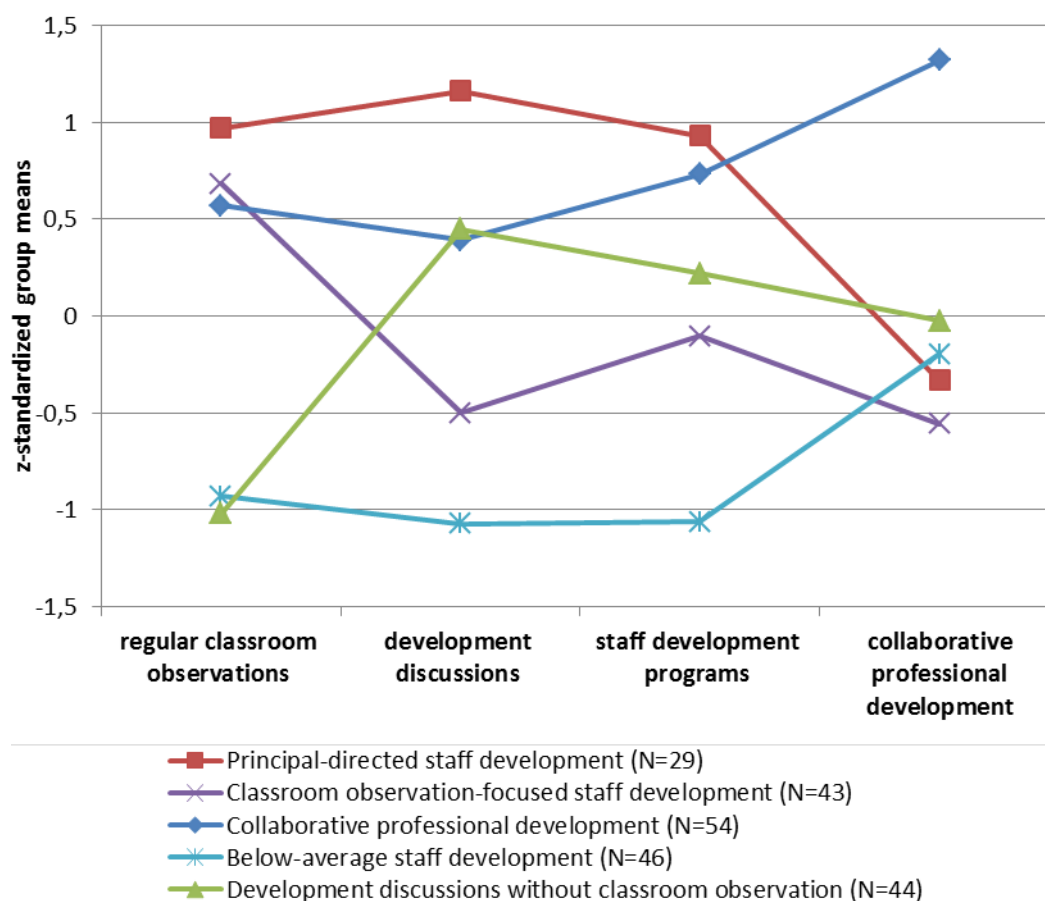
10

Figure 2. Staff development patterns – results of cluster analysis

The staff development patterns that could thus be identified are listed in the following:

(1) Principal-directed staff development: this pattern is characterized by an over-proportionate use of principal-directed staff development instruments (classroom observations, development discussions, staff development programs). Forms of collaborative professional development like student surveys or peer classroom observations were under-represented. There were 29 schools (13.4 percent) that could be assigned to this pattern.

(2) Collaborative professional development: a quarter of the schools (*n*=54) used this pattern. At these schools, staff development was predominantly characterized by forms of collaborative professional development activities, which were over-proportionately represented in comparison to other schools. Consequently, staff development in this case can be seen as mainly self-directed. However, simultaneously the three instruments of systematic staff development were also in use (slightly above-average). This indicates that the professionalization of staff is conducted systematically (up to a certain degree), whereby collaborative professional development activities may be part of a wider systematic staff development concept.

(3) Development discussions without classroom observation: this pattern was used by 44 schools (20.4 percent). It is characterized by an above-average use of development discussions. Furthermore, there is an average use of strategic staff development programs and collaborative professional development activities. Classroom observations were highly under-represented in this pattern. Principals who chose this pattern generally seemed to have a great interest in the professional development of their teachers. Nonetheless, it remains unclear which information they use (e.g. self-report of teachers, evaluation data) to plan professional development activities.

11

(4) Classroom observation-focussed staff development: this pattern is characterized by an above-average use of regular classroom observations by principals. There were 43 schools (19.9 percent) that could be assigned to this pattern. However, results of these observations did not seem to be used for development discussions or for systematic staff development programs, since these development instruments were not often applied. Forms of collaborative professionalization were also rarely found.

(5) Below-average staff development: in 46 schools (21.3 percent) staff development was generally sub-standard. However, forms of collaborative professionalization were over-represented.

(6) With respect to the association between patterns of staff development and context factors we can state that there were medium-sized effects (ω^2=0.16, p=0.000) between the type of school and the school-based staff development patterns used: primary schools often practiced a collaborative professional development, while systematic staff development in the school type Gymnasium, if practiced at all, was mostly limited to development discussions. A comparably large part of the other secondary schools practiced a classroom observation-focussed staff development (see Figure 3).
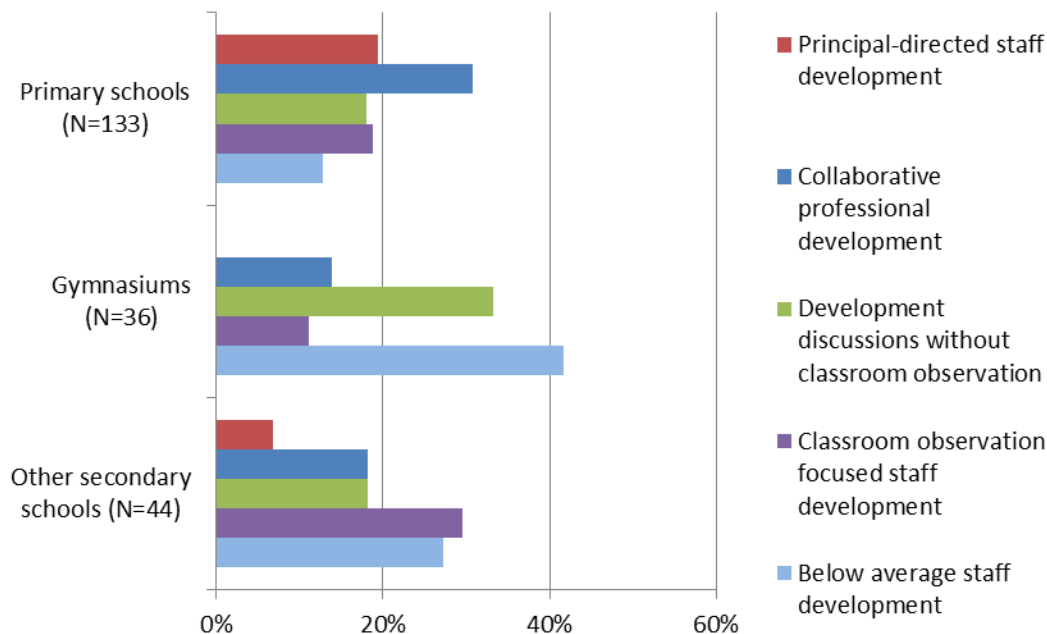

Figure 3. Patterns of staff development in different types of school

There were also medium-sized effects with respect to the federal state (ω^2=0.21, $p \leqslant 0.001$), the proportion of immigrant students (η^2=0.11, $p \leqslant 0.001$), as well as the age of the principal (η^2=0.08, $p \leqslant 0.05$). In Berlin, the proportion of schools with a below-average staff development was comparably high, while in Brandenburg the principal-directed staff development pattern was over-represented. In schools that featured below-average staff development, the proportion of immigrant students was on average highest, whereas in schools with a principal-directed staff development pattern it was lowest. Regarding the age of the principals, results showed that younger principals favored a classroom observation-focussed staff development, while older principals generally promoted little staff development (below-average staff development). There were small effects with regard to the school size (η^2=0.05, $p \leqslant 0.05$). A collaborative professional development pattern was rather used in small schools, while large schools generally employed less staff development instruments (below-average staff development). The age of the teaching staff did not have any significant effect.

(7) Regarding the relationship between the staff development patterns and the professional development activities of teachers, we found large differences between the different staff development patterns and the extent of professional development activities of teachers associated with each pattern (η^2=0.19, $p \leqslant 0.001$). As could be expected, schools featuring a below-average staff development also showed the lowest average extent of professional development activities of teachers. In schools with a principal-directed staff development pattern and schools featuring collaborative professional development, on average teachers showed the highest amount of professional development activities. Consequently, both patterns seem to be of significant importance for the extent of professional development activities of teachers. Furthermore, the classroom observation-focussed staff development seemed to be of relatively high importance for the extent of professional development activities of teachers. Schools in this cluster differed only slightly with regard to their average amount of teachers' professional development activities from schools with a principal-directed staff development pattern (d=0.42, $p \leqslant 0.01$) and schools with collaborative professional development (d=0.42, $p \leqslant 0.001$). The pattern development discussion without classroom observations has no significant impact on the professional activities.

6. Discussion and future directions

The present study comprehensively assessed staff development in two German states and examined impacts of different staff development instruments and patterns on the professional development activities of teachers.

To begin with, the descriptive results validated existing findings, confirming that staff development instruments are used very differently by principals. Relatively encouraging were the results regarding the use of classroom observations and development discussions. According to the majority of principals, both instruments were used regularly. Considering that principals are directly responsible for those instruments, their responses might have been affected by social desirability. Taking findings from Meetz (2007, p. 154) into account, according to which principals and teaching staff have very different perceptions of staff development measures, an additional survey of teachers would be advisable in future studies. Our results regarding the prevalence of collaborative professional development confirm previous findings in this area in that this form of professionalization is rarely established in German schools.

With regard to the extent of professional development activities of teachers, different staff development instruments have been proven to have a different impact. Staff development programs and regular classroom observations thereby showed the greatest correlations with developmental activities. It is not surprising that staff development programs have an impact on teachers' participation in professional development activities; however, what effect staff development programs have on the exact types of professional development activities remains unclear. The significance of classroom observations for the professional development activities of teachers might indicate that principals use these as an effective instrument that, in comparison to collaborative professional development activities (which is generally of little importance), is relatively easily incorporated into the practice of wider school improvement plans. More importantly, classroom observations present an opportunity for principals to inform themselves about the teaching of their staff, in order to make evidence-based decisions about, for instance, suitable professional development activities. Furthermore, it is conceivable thatteachers regard classroom observations as principals taking an interest in their work. The results of Blase and Blase (1998, pp. 109-111) support this hypothesis, since they found that the wandering around of principals had a positive effect on staff morale and the motivation of teachers. More specific research is needed to address the question of which specific

instruments (ratings) and procedures (feedback) of performance reviews principals should use in order to support the professional development of teachers.

The fact that the predictors in our model explained a third of the variance points to the importance of other factors, not included in this study, which increase the effectivity of staff development instruments. The review by Sleegers et al. (2002), for example, identified a high participation of teachers in decision-making processes and a collaborative culture among teachers as the main prerequisites for professional development. However, the exact influence teacher variables (e.g. self-efficacy, see Geijsel et al., 2009; or motivation, see Rzejak et al., 2014; Sleegers et al., 2014) exert over professional development activities still needs to be examined. In order to research the effectiveness of staff development measures adequately, longitudinal studies, similar to the recent one by Sleegers et al. (2014), are needed. In this study the authors explored the influence of a range of variables (e.g. leadership practices, school organizational conditions) on the classroom practices of teachers over a four year period, finding that organizational as well as teacher-level conditions (e.g. motivation) were relevant.

The identification of different staff development patterns demonstrated that the four instruments are used differentially in practice. While two patterns (i.e. principal-directed staff development and collaborative professional development) feature a very comprehensive use of a range of instruments, and a third pattern is characterized by a below-average use of all instruments (below-average staff development), there are also patterns that feature predominantly one instrument (i.e. development discussions without classroom observation and classroom observation-focussed staff development patterns). Furthermore, the use of the different staff development patterns seems to be heavily dependent on the type of school. Compared to secondary schools, there is a significantly higher prevalence of principal-directed staff development and collaborative professional development patterns in primary schools. Particularly principals of Gymnasiums were relatively passive with regard to their staff development. This can be explained by the fact that in Gymnasiums, in contrast to other school types, there are heads of the departments to whom subject-related development tasks can be delegated to. Follow-up studies would need to examine if that is indeed the case in Gymnasiums of this sample.

The finding that the schools with the below-average staff development pattern feature a relatively high proportion of immigrant students, might be due to a lack of resources for staff development in schools which face challenging circumstances. This is problematic because especially challenged schools are in need of effective instruments of staff development. That older principals engage in less staff development might be explained by a lower readiness for innovation of older principals (Moolenaar et al., 2010, p. 646) and possibly by fewer professional development activities they participated in that focussed on staff development and school improvement (Thillmann et al., 2015, p. 226).

With regard to the impact of different patterns, it could be shown that the two patterns with the highest use of staff development instruments also seemed to have the greatest effect on the professional development activities of teachers. Interestingly, however, one of the more "parsimonious" patterns of regular classroom observations was also demonstrated to be comparably effective with regard to the professional development activities of teachers. The fact that the pattern development discussions without classroom observation was not related to the professional development activities of the teachers indicates that development discussions are particularly useful when combined with classroom observations. This is most likely because the development discussions can then be based on information about the classroom practices elicited from the observation.

The current study is the first to examine the use of different instruments of school-based staff development comprehensively. Our results regarding the different patterns provide a first indication of which staff development strategies principals actually employ in practice. In addition to a replication of the identified patterns with larger samples that also takes the perspective of teachers into account,

further qualitative research is needed. Using qualitative methods (e.g. interviews, document analysis), it would further need to be demonstrated if and to what degree the use of different staff development instruments is in fact systematically interlinked. Thus, it would be interesting, for instance, to determine if and how results from classroom observations are used in practice as a basis for development discussions, performance targets, or professional development planning. Finally, further research would be needed that examines if the three most relevant staff development patterns identified in this study can also be proven to be effective with regard to somewhat "harder" criteria (like teachers' teaching skills or student achievement) than the extent of professional development activities of teachers. This would also be fruitful because our study did not determine if the professional development activities of teachers led to a qualitative improvement in teaching, which current studies suggest (e.g. Darling-Hammond et al., 2009).

Furthermore, we have to point out that the interpretation of our findings is limited by our use of a cross-sectional design. Consequently, complex causal relationships, which would need longitudinal designs, could not be examined. The data set in the present study was collected as part of a wider study named "Evaluation policies and practices in schools," funded by the Federal Ministry of Education and Research, which aimed at examining the enhanced autonomy in schools under the new governance model. Since the study did not focus on staff development, future research should aim to increase reliability by improving the quality of the scales we used and by assessing the constructs in a more complex manner.

In order to increase the generalizability of our findings, additional federal states would have to be included in the analysis. Due to the federal system in Germany, the states feature different legal structures and requirements that also extend to the degree of autonomy with regard to staff development. Our results regarding the prediction of professional development activities also point in this direction: the predictor federal state showed a significant influence on the dependent variable. It could for instance be explored if the here identified patterns of staff development can be replicated in other states. The current study allows for some first practical implications at least for the states of Berlin and Brandenburg. The high significance of classroom observations and staff development programs indicates that there are suitable instruments of school improvement available to principals. This finding could justify a more systematic incorporation of these measures into the relevant regulations (similarly to how Brandenburg has since incorporated development discussions, see Ministerium für Bildung, Jugend und Sport, 2012). Furthermore, it seems valuable to regard different instruments of staff development as well as their interactions, since this approach revealed, for example, that the impact of development discussions could be improved by preceding classroom observations.

Despite the methodological limitations, the current study is the first to examine staff development in German schools systematically. Our results provide some good leads for further studies in this area.

Notes

1. Since the 2011/2012 school year, all Berlin teachers have to evaluate their classes using the self-evaluation portal developed by the *Berlin-Brandenburg Institute for School Improvement* (Senatsverwaltung für Bildung, Wissenschaft und Forschung, 2011, p. 3).
2. In addition to the federal state of Baden-Württemberg, the study by Appius et al. (2012) also takes two Swiss cantons into account.
3. The main purpose of the pilot project "Independent School" was to foster school improvement through an increased autonomy and greater responsibility of the school, for instance with regard to staff development (see Holtappels et al., 2008).
4. Compared to other German secondary school types, the "Gymnasium" is characterized by a stronger academic focus. At the Gymnasium, students are being prepared to enter university.

5. In order to increase variance for the cluster analysis, we included the dichotomized variable "regular classroom observations" as an ordinal variable with three response categories (no regular classroom observations – once per school term – more often).

6. Data from principals were only included in the analysis if there were no missing data relevant to the analysis. The sample for the regression analysis was $n=139$ principals. Data from $n=216$ principals were included in the cluster analysis. Both samples did not show significant differences from the total sample with regard to the variables type of school and region.

7. According to Cohen (1988, p. 284) $\eta^2=0.01$ equals a small, $\eta^2=0.06$ a medium, and $\eta^2=0.14$ a large effect.

8. According to Cohen (1988, p. 40) $d=0.20$ equals a small, $d=0.50$ a medium, and $d=0.80$ a large effect.

9. Whereby $\omega^2=0.01$ constitutes a small, $\omega^2=0.09$ a medium, and $\omega^2=0.25$ a large effect (Rasch et al., 2010, p. 181).

10. M=arithmetic mean; SD=standard deviation.

11. These are two items from the scale "development discussions", which were assessed using a six-point scale (1=not true at all; 6=exactly true).

References

Appius, S., Steger Vogt, E., Kansteiner-Schänzlin, K. and Bach-Blattner, T. (2012), "Personalentwicklung an Schulen. Eine Bestandsaufnahme aus Sicht deutscher und schweizerischer Schulleitungen", *Empirische Pädagogik*, Vol. 26 No. 1, pp. 123-141.

Avenarius, H., Brauckmann, S., Döbert, H., Isermann, K., Kimmig, T. and Seeber, S. (2006), Durch größere Eigenverantwortlichkeit zu besseren Schulen. Ergebnisbericht der wissenschaftlichen Begleitung des "Modellvorhabens eigenverantwortliche Schule (MeS)" im Land Berlin, Deutsches Institut für Internationale Pädagogische Forschung (DIPF), Berlin and Frankfurt am Main.

Becker, M. (2009), Personalentwicklung, Schäffer-Poeschel, Stuttgart.

Blase, J. and Blase, J. (1998), Handbook of Instructional Leadership: How Really Good Principals Promote Teaching and Learning, Corwin Press, Thousand Oaks, CA.

Bonsen, M., Gathen, J., Iglhaut, C. and Pfeiffer, H. (2002), Die Wirksamkeit von Schulleitung – Empirische Annäherungen an ein Gesamtmodell schulischen Leitungshandelns, Juventa, Weinheim and München.

Bortz, J. (2005), Statistik für Human- und Sozialwissenschaftler, Springer Medizin Verlag, Heidelberg.

Brooks, J.S., Jean-Marie, G., Normore, A.H. and Hodgins, D.W. (2007), "Distributed leadership for social justice: exploring how influence and equity are stretched over an urban high school", Journal of School Leadership, Vol. 17 No. 4, pp. 378-408.

Buhren, C.G. and Rolff, H.-G. (2009), Personalmanagement für die Schule. Ein Handbuch für Schulleitung und Kollegium, Beltz, Weinheim/Basel.

Burkard, C. (2000), "Fortbildungsplanung – ein instrument zur Qualitätsentwicklung von Schulen", SchulVerwaltung spezial, Special Issue 3, pp. 19-21.

Cohen, J. (1988) Statistical Power Analysis for the Behavioral Sciences, Lawrence Erlbaum Associates, Hillsdale, NJ.

Craft, A. (2000), Continuing Professional Development. A Practical Guide for Teachers and Schools, RoutledgeFalmer, London.

Darling-Hammond, L., Chung Wei, R., Andree, A. and Orphanos, S. (2009), Professional Learning in the Learning Profession: A Status Report on Teacher Development in the United States and Abroad, National Staff Development Council, Stanford University, Stanford, CA.

David, J. (2008), "Classroom walkthroughs", Educational Leadership, Vol. 65 No. 4, pp. 81-82.

Desimone, L., Porter, A.C., Garet, M., Suk Yoon, K. and Birman, B. (2002), "Effects of professional development on teachers' instruction: results from a three-year longitudinal study", Educational Evaluation and Policy Analysis, Vol. 24 No. 2, pp. 81-112.

Desimone, L.M. (2009), "Improving impact studies of teachers' professional development: toward better conceptualizations and measures", Educational Researcher, Vol. 38 No. 3, pp. 181-199.

Downey, C.J., Steffy, B.E., Poston, W.K. Jr and English, F.W. (2010), Advancing the Three-Minute walk-Through: Mastering Reflective Practice, Corwin Press, Thousand Oaks, CA.

Dubs, R. (2005), Die Führung einer Schule. Leadership und Management, Franz Steiner Verlag, Wiesbaden.

Feldhoff, T., Kanders, M. and Rolff, H.-G. (2008), "Qualitätssicherung und Rechenschaftslegung", in Holtappels, H.G., Klemm, K. and Rolff, H.-G. (Eds), Schulentwicklung durch Gestaltungsautonomie. Ergebnisse der Begleitforschung zum Modellvorhaben, Selbstständige Schule' in Nordrhein-Westfalen, Waxmann, Münster, pp. 183-194.

Franke, M.L., Carpenter, T.P., Levi, L. and Fennema, E. (2001), "Capturing teachers' generative change: a follow-up study of professional development in mathematics", American Educational Research Journal, Vol. 38 No. 3, pp. 653-689.

Geijsel, F., Sleegers, P., Stoel, R. and Krüger, M. (2009), "The effect of teacher psychological, school organizational and leadership factors on teachers' professional learning in Dutch schools", Elementary School Journal, Vol. 109 No. 4, pp. 406-427.

Goldman, E. (1998), "The significance of leadership style", Educational Leadership, Vol. 55 No. 7, pp. 20-22.

Goldschmidt, P. and Phelps, G. (2010), "Does teacher professional development affect content and pedagogical knowledge: how much and for how long?", Economics of Education Review, Vol. 29 No. 3, pp. 432-439.

Gräsel, C., Fussangel, K. and Parchmann, I. (2006), "Lerngemeinschaften in der Lehrerfortbildung. Kooperationserfahrungen und -überzeugungen von Lehrkräften", Zeitschrift für Erziehungswissenschaft, Vol. 9 No. 4, pp. 545-561.

Guskey, T.R. (2000), Evaluating Professional Development, Corwin Press, Thousand Oaks, CA.

Holling, H. and Liepmann, D. (2007), "Personalentwicklung", in Schuler, H., Brandstätter, H.,

Bungard, W., Greif, S., Ulich, E. and Wilpert, B. (Eds), Lehrbuch Organisationspsychologie, Huber, Bern, pp. 345-383.

Holtappels, H.G. and Müller, S. (2002), "Inhalte und Struktur von Schulprogrammen. Inhaltsanalyse der Schulprogrammtexte Hamburger Schulen", in Rolff, H.-G., Holtappels, H.-G., Klemm, K., Pfeifer, H. and Schulz-Zander, R. (Eds), Jahrbuch der Schulentwicklung. Daten, Beispiele und Perspektiven (Band 12), Juventa, Weinheim, pp. 209-231.

Holtappels, H.G., Klemm, K. and Rolff, H.-G. (Eds) (2008), Schulentwicklung durch Gestaltungsautonomie. Ergebnisse der Begleitforschung zum Modellvorhaben, 'Selbstständige Schule' in Nordrhein-Westfalen, Waxmann, Münster.

Huber, S.G. (2013), Handbuch Führungskräfteentwicklung. Grundlagen und Handreichungen zur Qualifizierung und Personalentwicklung im Schulsystem, Carl Link/Wolters Kluwer Deutschland, Köln.

Kachur, D.S., Stout, J.A. and Edwards, C.L. (2010), Classroom Walkthroughs to Improve Teaching and Learning, Eye on Education Inc, Larchmont, NY.

Kieser, A. and Kubicek, H. (1983), Organisation, 2nd ed., de Gruyter, Berlin.

Klemm, K. and Meetz, F. (2008), "Personalmanagement und Sachmittelbewirtschaftung", in Holtappels, H.G., Klemm, K. and Rolff, H.-G. (Eds), Schulentwicklung durch Gestaltungsautonomie, Waxmann, Münster, pp. 174-182.

Leithwood, K., Louis, K., Anderson, S. and Wahlstrom, K. (2004), Review of Research. How Leadership Influences Learning, The Wallace Foundation, New York, NY.

Lipowsky, F. (2010), "Lernen im Beruf. Empirische Befunde zur Wirksamkeit von Lehrerfortbildung", in Müller, F., Eichenberger, A., Lüders, M. and Mayr, J. (Eds), Lehrerinnen und Lehrer lernen – Konzepte und Befunde zur Lehrerfortbildung, Waxmann, Münster, pp. 51-72.

Little, R.J.A. (1988), "Missing data adjustments in large surveys", Journal of Business Economics and Statistics, Vol. 6 No. 3, pp. 287-301.

Ministerium für Bildung, Jugend und Sport (2012), "Verwaltungsvorschriften über die Führung eines Leistungs-und Entwicklungsgesprächs mit Lehrkräften an öffentlichen Schulen (VV-Leistungs – und Entwicklungsgespräch-Lehrkräfte – VVLEG-L)", Amtsblatt des Ministeriums für Bildung, Jugend und Sport, Vol. 21 No. 7, pp. 286-290, available at: www.bravors.brandenburg.de/sixcms/media.php/15/Abl-MBJS_07_2012.pdf (accessed December 1, 2014).

Meetz, F. (2007), Personalentwicklung als Element der Schulentwicklung. Bestandsaufnahme und Perspektiven, Klinkhardt, Bad Heilbrunn.

Mintzberg, H. (1983/1992), Die Mintzberg-Struktur: Organisation effektiver gestalten, (Trans. by H. Höhlein), Verlag Moderne Industrie, Landsberg/Lech.

Mohr, I. (2006), Analyse von Schulprogrammen. Eine Arbeit im Rahmen der Internationalen Grundschul-Lese-Untersuchung (IGLU), Waxmann, Münster.

Moolenaar, N.M., Daly, A.J. and Sleegers, P.J.C. (2010), "Occupying the principal position: examining relationships between transformational leadership, social network position, and schools' innovative climate", Educational Administration Quarterly, Vol. 46 No. 5, pp. 623-670.

Ngo, H.-Y., Turban, D., Lau, C.-M. and Lui, S.-Y. (1998), "Human resource practices and firm performance of multinational corporations: influences of country origin", The International Journal of Human Resource Management, Vol. 9 No. 4, pp. 632-652.

Ovando, M. and Ramirez, A. (2007), "Principals' instructional leadership within a teacher performance appraisal system: enhancing students' academic success", Journal of Personnel Evaluation in Education, Vol. 20 Nos 1-2, pp. 85-110.

Parr, J., Timperley, P., Reddish, R., Jesson, R. and Adams, R. (2007), Literacy Professional Development Project: Identifying Effective Teaching and Professional Development Practices for Enhanced Student Learning, Ministry of Education, Wellington.

Rasch, B., Friese, M., Hofmann, W. and Naumann, E. (2010), Quantitative Methoden 2. Einführung in die Statistik, Springer Medizin Verlag, Heidelberg.

Robinson, V.M., Lloyd, C.A. and Rowe, K.J. (2008), "The impact of leadership on student outcomes: an analysis of the differential effects of leadership types", Educational Administration Quarterly, Vol. 44 No. 5, pp. 635-674.

Rosenbusch, H.S. and Huber, S.G. (2001), "Qualifizierungsmaßnahmen von Schulleiterinnen und Schulleitern in den Ländern der Bundesrepublik Deutschland", Schulmanagement, Vol. 32 No. 4, pp. 8-16.

Rosenbusch, H.S., Braun-Bau, S. and Warwas, J. (2006), Schulleitungtätigkeit an bayerischen Grund-, Haupt- und Realschulen. Bestandsaufnahme und Vorschläge für eine Neuorientierung. Forschungsstelle für Schulentwicklung und Schulmanagement der Universität Bamberg, Universitätsdruck, Bamberg.

Rzejak, D., Künsting, J., Lipowsky, F., Fischer, E., Dezhgahi, U. and Reichardt, A. (2014), "Facetten der Lehrerfortbildungsmotivation – eine faktorenanalytische Betrachtung", Journal for educational research online, Vol. 6 No. 1, pp. 139-159.

Senatsverwaltung für Bildung, Wissenschaft und Forschung (2011), "Qualitätspaket Kita und Schule", available at:
www.berlin.de/imperia/md/content/sen-
bildung/schulqualitaet/qualitaetspaket.pdf?start&ts=1304676467&file=qualitaetspaket.pdf
(accessed December 1, 2014).

Senatsverwaltung für Bildung, Wissenschaft und Forschung (2013), "Handlungsrahmen Schulqualität in Berlin. Qualitätsbereiche und Qualitätsmerkmale", available at:
www.berlin.de/imperia/md/content/sen-
bildung/schulqualitaet/handlungsrahmen_schulqualitaet.pdf?start&ts=1383302071&file=handl
ungsrahmen_schulqualitaet.pdf (accessed December 1, 2014).

Senkbeil, M., Drechsel, B., Rolff, H.-G., Bonsen, M., Zimmer, K., Lehmann, R.H. and Neumann, A. (2004), "Merkmale und Wahrnehmungen von Schule und Unterricht", in Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M., Pekrun, R., Rolff, H.-G., Rost, J. and Schiefele, U. (Eds), PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs, Waxmann, Münster, pp. 296-314.

Sleegers, P., Geijsel, F. and van den Berg, R. (2002), "Conditions fostering educational change", in Leithwood, K. and Hallinger, P. (Eds), Second International Handbook of Educational Leadership and Administration, Kluwer Academic Publishers, Dordrecht, pp. 75-102.

Sleegers, P., Thoonen, E., Oort, F. and Peetsma, T. (2014), "Changing classroom practices: the role of school-wide capacity for sustainable improvement", Journal of Educational Administration, Vol. 52 No. 5, pp. 617-652.

Smith, C. and Gillespie, M. (2007), "Research on professional development and teacher change: implications for adult basic education", in Comings, J., Garner, B. and Smith, C. (Eds), Review of Adult Learning and Literacy, Vol. 7, Lawrence Erlbaum Associates, Mahwah, NJ, pp. 205-244.

Smither, J., London, M. and Reilly, R. (2005), "Does performance improve following multisource feedback? A theoretical model, meta-analysis, and review of empirical findings", Personnel Psychology, Vol. 58 No. 1, pp. 33-66.

Sonntag, K. and Stegmaier, R. (2006), "Personalentwicklung und Unternehmensperformance – Eine Evaluationsperspektive für das human resource management", in Sonntag, K. (Ed), Personalentwicklung in Organisationen, 3rd ed., Hogrefe, Göttingen, pp. 389-410.

Spillane, J., Halverson, R. and Diamond, J. (2004), "Towards a theory of school leadership practice: implications of a distributed perspective", Journal of Curriculum Studies, Vol. 36 No. 1, pp. 3-34.

Steger Vogt, E. (2013), Personalentwicklung – Führungsaufgabe von Schulleitungen, Waxmann, Münster.

Stoll, L. and Louis, S. (2007), Professional Learning Communities: Divergence, Detail and Difficulties, Open University Press, Maidenhead.

Terhart, E. (2010), "Personalauswahl, Personaleinsatz und Personalentwicklung an Schulen", in Altrichter, H. and Maag Merki, K. (Eds), Handbuch neue Steuerung im Schulsystem, VS Verl. für Sozialwissenschaften, Wiesbaden, pp. 255-275.

Thillmann, K. (2012), Schulentwicklung und Schulorganisation. Eine empirische Untersuchung schulischer Organisationsgestaltung vor dem Hintergrund der Neuen Steuerung im Bildungssystem, Freie Universität Berlin, Berlin.

Thillmann, K., Brauckmann, S., Herrmann, C. and Thiel, F. (2015), "Praxis schulischer Personalentwicklung unter den Bedingungen der Neuen Steuerung", in Abt, H.J., Brüsemeister, T., Schemmann, M. and Wissinger, J. (Eds), Governance im Bildungssystem,

Analysen zur Mehrebenenperspektive, Steuerung und Koordination, Educational Governance, Vol. 26, Springer VS, Wiesbaden, pp. 195-228, available at: http://dx.doi.org/10.1007/978-3-658-06523-2_8 (accessed December 1, 2014).

Tuytens, M. and Devos, G. (2011), "Stimulating professional learning through teacher evaluation: an impossible task for the principal?", Teaching and Teacher Education, Vol. 27 No. 5, pp. 891-899.

Vereinigung der Bayerischen Wirtschaft e.V. (2010), Bildungsautonomie: Zwischen Regulierung und Eigenverantwortung – die Bundesländer im Vergleich. Expertenrating der Schul – und Hochschulgesetze der Länder zum Jahresgutachten 2010, Verlag für Sozialwissenschaften, München.

Waters, T., Marzano, R.J. and McNulty, B. (2003), "Balanced leadership: what 30 years of research tells us about the effect of leadership on pupil achievement", working paper, Mid-continent Research for Education and Learning (McREL), Denver, CO.

Wissinger, J. (2002), "Schulleitung im internationalen Vergleich – Ergebnisse der TIMSS-Schulleiterbefragung", in Wissinger, J. and Huber, S.G. (Eds), Schulleitung - Forschung und Qualifizierung, Leske and Budrich, Opladen, pp. 45-61.