Frontline Learning Research
An official journal of EARLI

Frontline **L**earning **R**esearch Vol. 3 No. 2 (2015) 27 - 46
ISSN 2295-3159

What Works in School? Expert and Novice Teachers' Beliefs about School Effectiveness

Johanna Fleckenstein[a], Friederike Zimmermann[b], Olaf Köller[a], Jens Möller[b]

[a]Leibniz Institute for Science and Mathematics Education, Germany

[b]Kiel University, Germany

Article received 4 April 2015 / revised 4 April 2015 / accepted 27 April 2015 / available online 4 June 2015

Abstract

In 2009, John Hattie first published his extensive metasynthesis concerning determinants of student achievement. It provides an answer to the question: "What works in school?" The present study examines how this question is answered by pre- and in-service teachers, how their beliefs correspond to the current state of research and whether they differ according to the teachers' level of expertise. Thus, it takes on a novel approach as it draws on data from two sources in the field of education -- empirical research and teachers' beliefs -- and examines their similarities and differences. The teachers' beliefs were elicited by asking N = 729 participants to estimate the effect sizes of several determinants of student achievement. Those were compared to the empirical effect sizes found by Hattie (2009). Profile correlations showed that expert teachers' beliefs are more congruent with current research findings than those of novice teachers. We further examined where expert and novice teachers' beliefs differ substantially from each other by using confirmatory factor analysis (CFA) and comparing group means in latent variables. Our findings suggest that teachers' beliefs about school effectiveness are related to professional experience: Expert teachers showed a stronger overall congruence with empirical evidence, scoring higher in achievement-related variables and lower in variables concerning surface- and infrastructural conditions of schooling as well as student-internal factors. Results are discussed with regard to teacher-education practices that emphasize research findings and challenge existing beliefs of (prospective) teachers.

Keywords: Teacher beliefs; teacher education; professional competence; school effectiveness

Corresponding author: Johanna Fleckenstein, Leibniz Institute for Science and Mathematics Education, Department of Educational Research, Olshausenstr. 62, 24118, Kiel, Germany. Email address: fleckenstein@ipn.uni-kiel.de DOI: http://dx.doi.org/10.14786/flr.v3i2.162

Teachers' beliefs are often guided by subjective experience rather than by empirical data. Thus, it is to be expected that they generally diverge from research findings. This assumption also pertains to the specific case of teachers' beliefs that manifest in their response to the question: "What works in school?" School effectiveness research has tried to answer this question, the latest attempt being Hattie's (2009, 2012) metasynthesis of factors that influence student achievement. However, little is known of the practitioner's answer to one and the same question: What factors do (expert and novice) teachers believe to influence student achievement? Where do their beliefs differ from research findings substantially?

These questions are highly relevant as teachers' beliefs have been shown to influence teaching and learning. If they differ substantially from results of school effectiveness research, we have reason to assume a negative effect on educational outcomes. For example, if a teacher undervalues a particular teaching method or overvalues surface structural aspects like class size, this could be a serious threat to effective teaching. An investigation into teachers' beliefs concerning school effectiveness can show us where these discrepancies are and thus inform teacher education practice and classroom instruction. Teachers' beliefs are typically represented as part of a multi-dimensional construct of teachers' professional competence (Baumert & Kunter, 2006). They influence teachers' perceptions and judgments and, consequently, affect their classroom instruction (Calderhead, 1996; Pajares, 1992). Beliefs are subject to change and thus can be expected to differ in expert and novice teachers.

In the present study, we are particularly interested in a specific subset of teachers' beliefs, namely those regarding the effectiveness of school- and education-related factors. In the last few decades, there has been a lot of research concerning the question of what works in school – and what does not (Fraser, Walberg, Welch & Hattie, 1987; Walberg, 1986; Hattie, 2003; Wang, Haertel, & Walberg, 1990, 1993). The most recent and extensive example of this is Hattie's synthesis of meta-analyses (2009, 2012), in which he examines the influence of 138 factors on student achievement. Teachers should be familiar with such findings of school effectiveness research in order to make informed decisions and focus on the most effective interventions. For this kind of evidence-based practice, however, teachers not only have to *know* about such findings but actually to *believe* that they are true. Hence, we asked pre-service ("novice") and in-service ("expert") teachers for their beliefs about the efficacy of certain determinants of student achievement. These ratings of our expert and novice teachers were then compared with each other and contrasted with the findings of Hattie (2009).

In the following we will provide a brief theoretical background on teachers' beliefs. Since the body of research on teachers' beliefs is quite extensive we concentrate on the following relevant aspects: the theoretical construct of teachers' beliefs, the influence of teachers' beliefs on classroom processes and student outcomes, the issue of teachers' beliefs being guided by subjective experience rather than objective fact, and the general differences in beliefs of novice and expert teachers. Subsequently, we locate teachers' beliefs within a model of teachers' professional competence, which – in accordance with the prevailing *expert(-novice) paradigm* – suggests the malleability of all its components. The beliefs examined here focus on determinants of student achievement, therefore, we will also show the most important and recent results of school effectiveness research. As Hattie (2009) serves as the basis of our study, particular attention is paid to his comprehensive aggregation of existing meta-analyses (*metasynthesis*; see Zell & Krizan, 2014).

1. Theoretical Background

1.1 Teachers' beliefs

According to Barcelos (2003), beliefs are types of thoughts which provide a basis for decisions and actions. Harvey (1986) described a belief system as "a set of conceptual representations which signify to its

holder a reality or a given state of affairs of sufficient validity, truth or trustworthiness to warrant reliance upon it as a guide to personal thought and action" (p. 146). Beliefs in general are thought of as psychologically held understandings, premises, or propositions about the world that a person perceives as being veritable (Richardson, 1996). Teachers' beliefs can be seen as a substructure of the general belief system. They consist of beliefs that serve as a guide when dealing with school- and instruction-related situations. In educational settings, Haney et al. (2003) defined beliefs as "one's convictions, philosophy, tenets, or opinions about teaching and learning" (p. 367). As such, teachers' beliefs may include subjective theories about how students learn, what a teacher should or should not do and which instructional strategies work effectively.

The last few decades have brought out a substantial body of research on the beliefs of teachers (for comprehensive research reviews see Calderhead, 1996; Fang, 1996; Pajares, 1992; Nespor, 1987; Richardson, 1996; Stuart & Thurlow, 2000; Verloop et al., 2001; Wenden, 1999; Woods, 1996; Zheng, 2009). Teachers' beliefs influence perception and judgment which in turn guide their actions in the context of school and education (Pajares, 1992). Prior research has shown that teachers' beliefs have a critical impact on the way they teach in the classroom, learn how to teach, and perceive educational reforms (M. Borg, 2001; Allen, 2002; S. Borg, 2003; Freeman, 2002; Yook, 2010). Other studies have shown the importance of teachers' beliefs for student achievement (Peterson, Fennema, Carpenter & Loef, 1989; Staub & Stern, 2002). An especially well-researched issue in this context are teachers' self-efficacy beliefs. According to Tschannen-Moran et al. (1998, p. 233), teacher efficacy is "the teacher's belief in his or her capability to organize and execute courses of action required to successfully accomplish a specific task in a particular context". Such beliefs have been shown to critically influence a teacher's performance and motivation (Bandura, 1997; Ross,1998; Tschannen-Moran & Woolfolk Hoy, 2001; Tschannen-Moran, Woolfolk Hoy & Hoy,1998; Woolfolk & Hoy, 1990; Woolfolk, Rosoff & Hoy, 1990; Woolfolk Hoy & Davis, 2006) as well as his or her students' achievement in school (Bates, Latham & Kim, 2011; Muijs & Rejnolds, 2001; Ross, 1992, 1998).

Against this background it becomes evident that the beliefs of teachers are an important issue for teaching and learning. The definitions above show that beliefs are always perceived to be correct by the individual. However, this is not necessarily the case: Teachers' beliefs – as a subgroup of beliefs in general – are especially likely to be flawed in the sense that they contradict empirical evidence. In the following we discuss why this is and in what way it can lead to substantial problems, in particular for novice teachers.

Teachers' beliefs are highly subjective, tend to be persistent and develop at a rather early stage in life (Lortie, 1975; Pajares, 1992). This is partly due to the long-term experience with schools and classrooms during a teacher's own time as a student, which serves as the starting point of his or her training and career. Hattie (2009) calls this relatively stable system of beliefs the "grammar of schooling" (p. 5). It contains tacit and simplified notions of what a good teacher is and how students are supposed to behave (Clark, 1988; Nespor, 1987). As such, this belief system often diverges from empirical findings and can fatally influence the process of teaching (Kunter & Pohlmann, 2009). Understanding and challenging one's own beliefs is therefore considered an important aspect of teacher qualification (Bromme, 1997; Woolfolk-Hoy, Davis & Pape, 2006). However, this does not seem to be an easy task, since not even the confrontation with dissonance (e.g., induced by empirical evidence that contrasts one's beliefs) necessarily leads to a corresponding change in beliefs (Hart, 2004; Pajares, 1992).

Hattie (2009) claims that teachers experience almost everything they do in the classroom to have a positive influence on their students' learning. The remarkable differences in effectiveness of their efforts are easily overlooked due to the lack of firsthand comparison, since teachers are usually confined to their own classroom. They see that what they do seems to work fine, as (almost) everything a teacher does leads to an increase in students' achievement. Thus, there is always anecdotal evidence for the effectiveness of certain methods, even though the actual extent to which they support students' learning often differs dramatically. Hattie describes this basic principle of teaching as "just leave me alone as I have evidence that what I do enhances learning and achievement" (Hattie, 2009, p. 6). Admittedly, this is a rather simplified account of

teachers' self-efficacy beliefs. Most teachers indeed struggle with the teaching methods they use, try out new things and reflect whether or not they seem to be working. However, they often do this within their own frame of reference, not based on research evidence.

The fact that teachers' beliefs do not necessarily coincide with empirical findings can be problematic if not reflected thoroughly. Especially novices can have inadequate notions of what constitutes good teaching. Weinstein (1989) found that on average, pre-service teachers overestimate affective and social variables of classroom instruction (such as *patience* and the *ability to relate to children*) and underestimate cognitive and academic variables (such as *organization* and *challenging*). However, when contrasted with the perspectives of educational policy makers and researchers, in-service teachers' beliefs seem very similar to those of their less experienced colleagues. While the former two speak of good teaching in terms of outcomes in standardized assessment and direct instruction models (policy makers), as well as 'masterful teachers' with a whole set of well-defined professional skills (researchers; e.g. Shulman, 1987), the latter two have a notion of 'good teachers' that can be described as "warm, caring individuals who enjoy working with children" (Weinstein, 1989, p. 59). Hogan, Rabinowitz, and Craven (2003) compared novice and expert teachers and found that student achievement was important for expert teachers, while novice teachers paid more attention to student interest.

1.2 Teachers' professional competence

Teacher training and professional development are central issues in the international discussion on teacher effectiveness (Bauer & Prenzel, 2012; Cochran-Smith & Zeichner, 2005; Darling-Hammond & Bransford, 2005). The development of standards in teacher education requires an explicit analysis of challenges teachers face in their everyday professional life. Moreover, it demands a specification of competences necessary to master these challenges. The quest for *the good teacher* is not new; however, the specific facets of teachers' professional competence and the premise that teachers' cognitions are modifiable by means of training are direct results of a relatively recent objective in teacher(-education) research: the expert(-novice) paradigm (Berliner, 2004; Bromme, 1997, 2003; Ericsson & Lehmann, 1996; Ericsson, Charness, Feltovich & Hoffman, 2006). Accordingly, this paper is based on the central assumptions that (a) good teachers are experts of learning and teaching, and (b) they achieve this expertise in the form of professional competence through continuous teacher education and professional experience.

Experts are roughly defined as "individuals who exhibit reproducibly superior performance on representative, authentic tasks in their field" (Ericsson, 2006, p. 688). It is assumed that teachers' expertise or professional competence is acquired throughout pre- and in-service training as well as by hands-on experience in the classroom (Berliner, 2004). Teachers' professional competence is usually represented as a multi-dimensional construct. As such, based on the five core propositions of the National Board for Professional Teaching Standards (NBPTS), Baumert and Kunter (2006) proposed a model of teachers' professional action competence with four non-hierarchical dimensions: (1) specific declarative and procedural knowledge, which further distinguishes between content knowledge (CK), pedagogical knowledge (PK), and pedagogical content knowledge (PCK) (Shulman, 1986, 1987); (2) professional beliefs, values, subjective theories, normative preferences and objectives; (3) motivational orientations; and (4) meta-cognitive skills and professional self-regulation. In line with the expert(-novice) paradigm we assume that all of these competencies are subject to change throughout a teacher's professional life.

The boundaries between these categories of teachers' cognition, however, are more or less fuzzy: Knowledge – PK and PCK in particular – and beliefs are strongly interrelated theoretical constructs, even though they rely on dissimilar epistemological notions, as „belief is based on evaluation and judgment; knowledge is based on objective fact" (Pajares, 1992, p. 313). One and the same response to a pedagogical question can either demonstrate well-founded knowledge or it can be based on subjective belief. The two answers may differ in their epistemological status, though their distinction – philosophically speaking – is a mere social construct: Leatham (2006) argues that beliefs (things we *just believe*) and knowledge (things we

more than believe) can be viewed as complementary subsets of the things we believe. In comparison to belief, knowledge is characterized by a higher degree of certainty, for example, by being grounded in empirical evidence. In many empirical studies on teacher beliefs, however, the distinction between knowledge and beliefs is rather blurry. It is very difficult to distinguish whether teachers refer to their knowledge or beliefs when they plan, make decisions, or act in classroom (Verloop, Van Driel, & Miejer, 2001). In the present study, we use the concept of teachers' beliefs to refer to cognitions of teachers that are subjective and normative in nature, while they may or may not coincide with the more objective construct of knowledge.

1.3 School effectiveness research

The present study deals with teachers' beliefs concerning factors of school effectiveness. Thus, in the following we briefly summarize the central findings of school effectiveness research. We focus on proximal vs. distal aspects of schooling as there is a broad consensus concerning this dichotomy in the literature. The analyses performed in this paper focus on the broad categories school, teaching and student, so we aim at summarizing the comprehensive literature in school effectiveness research on this rather general level. Subsequently, we give a more detailed overview of Hattie's 2009 metasynthesis, concentrating on those variables and categories that were used in our questionnaire to elicit teacher' beliefs.

In the last few decades, increased efforts were made in school effectiveness research to study the importance of a range of determinants on successful schooling. The question of what works in school (and what does not) has been the central issue of a number of meta-analyses and research syntheses (Coleman et al., 1966; Fraser et al., 1987; Hattie, 2003, 2009, 2012; Jencks et al., 1973; Scheerens & Bosker, 1997; Seidel & Shavelson, 2007; Walberg, 1986; Wang et al., 1990, 1993). Those studies do not always agree on the specific size of an effect, however, there is a general tendency with regards to certain factors of school effectiveness. In general, the majority of these studies suggest that the amount of variance explained by proximal – school- or classroom-related – variables is considerable, and has a greater influence on student learning than more distal aspects such as school system and educational policy (Seidel & Shavelson, 2007). The general emphasis on proximal variables was, for example, shown by Scheerens and Bosker (1997). They combined the results of three meta-analyses as well as a re-analysis of an international data set and found that school-organizational factors (e.g., monitoring/ evaluation, orderly climate), instructional conditions (e.g., opportunity to learn, homework), and aspects of structured teaching (e.g., feedback, cooperative learning) are a better explanation for the differences between the achievement of students than more distal aspects such as resource input factors (e.g., student-teacher ratio, teachers' salary).

Moreover, the rank-ordering presented by Wang et al. (1993) put student characteristics and classroom practices ahead of design of program and school demographics. They found particularly strong effects for the variables meta-cognition, classroom management and quantity of instruction. Fraser et al. (1987) found the highest correlations with performance tests for variables related to student characteristics (especially cognitive ones), learning strategies, and structured or direct teaching. The results also revealed that open teaching and individualization are less powerful factors, at least when the dependent variable is (cognitive) achievement. Similar findings were shown by Walberg (1986). There have been many attempts to find a comprehensive consensus with regards to the effects of certain factors of school effectiveness. Scheerens (2004) presented the effectiveness enhancing conditions of schooling in five review studies (Cotton, 1995; Levine & Lezotte, 1990; Purkey & Smith, 1983; Sammons, Hillman & Mortimore, 1995; Scheerens, 1992): A consensus was reached with respect to many instruction-related factors, such as achievement orientation, high expectations, frequent testing/ monitoring, professional development, and structured or purposeful teaching.

Hattie's (2009) synthesis of over 800 meta-analyses was one of the most recent milestones in school effectiveness research: 52,637 individual studies with over 83 million students were used in order to determine the relevance of 138 factors for student achievement. For each of these factors he determined

Cohen's *d* as the averaged effect size. As a convention in the school context, effect sizes of $d > .40$ are considered substantial, since this would imply greater effects than one year of average schooling (Köller, 2012); Hattie calls this the *zone of desired effects*. Hence, the point of reference for the effectiveness of an innovation is not $d = 0$, but $d = .40$.

Hattie's results were largely in line with the prior findings in school effectiveness research as described in the preceding paragraphs. He systematized the individual factors according to six superordinate categories: *student*, *teacher*, *teaching*, *curriculum*, *school*, and *family*. The central results can be summarized as follows: More or less ineffective factors with $d < .40$ were primarily infrastructural conditions of schooling, such as within- or between-class grouping, finances, and reduction of class size. Moreover, aspects of the surface structure of teaching, which is often associated with progressive teaching approaches (e.g. open learning, multi-grade/-age classes, team teaching), did not show to be very effective either. These results may be surprising considering the socio-political discourse on education; however, against the background of modern classroom research they are to be expected: Research has shown that successful learning can be better predicted by the deep structure of teaching and learning than by the surface structure (e.g., Seidel & Shavelson, 2007). The latter can be observed and described without much effort, while the former requires more elaborate assessment. The use of surface-structure learning methods is not beneficial by itself, but only if it affects the level of deep-structure cognitive processing (e.g., by giving constructive feedback or teaching meta-cognitive strategies). In line with prior research on school effectiveness, high effect sizes could also be shown for cognitive and emotional student characteristics (e.g. prior knowledge, motivation) and instructional, achievement-related variables such as direct instruction and high expectations of the teacher. In agreement with prior research, Hattie's findings suggest that more distal factors are less important than proximal factors and that the structural conditions of teaching are less important than the process of teaching itself. The results also highlight the importance of the students' cognitive and non-cognitive prerequisites for learning.

2. The present study

In our study we attempted a direct comparison of the results of school effectiveness research (i.e. the effect sizes from Hattie's study) with the beliefs of novice and expert teachers (i.e. their ratings of effect sizes). This was a rather novel approach; however, Wang et al. (1993) adopted a similar strategy when they compared the results of 91 meta-analyses with the ratings of experts in education, namely 61 distinguished educational researchers. The correlation they found between expert ratings and meta-analyses was .59 ($p < .01$). The authors concluded that there is a general agreement between expert ratings and the meta-analyses regarding the effect of different variables on student learning and their relative strength. While Wang et al. (1993) examined the judgments of experts in educational research, our study dealt with the beliefs of teachers that are either enrolled in a teacher training program or work as teachers and school administrators. The objective of Wang et al. was to build a knowledge base in school effectiveness research: They used three different methods – content analyses, expert ratings, and results from meta-analyses – to quantify the importance and consistency of variables that influence student learning. Our objective, on the other hand, was to explicitly address the beliefs of those groups that actually are or are going to be working in the field and directly influence classroom processes. Prior research on teachers' beliefs (in general and especially those of novice teachers) has shown that they tend to be very subjective and are rather unlikely to be guided by empirical evidence, so we had reason to assume that is also the case also for beliefs concerning determinants of student achievement. Thus, we expected our pre- and in-service teachers' beliefs to differ more strongly from the findings of school effectiveness research than the ratings of expert researchers. With the epistemological question in mind that we raised above, one could also argue that Wang et al. (1993) examined knowledge while we examined beliefs.

School effectiveness research gives us a good theoretical understanding of what works in school. However, we have reason to assume that teachers' cognitions are not congruent with these findings, since their beliefs are at risk to be guided by subjective experience and beliefs rather than by empirical data. The present study focused on teachers' beliefs about which factors determine their students' achievement. Hence, the central questions were:

a) What are teachers' beliefs about the impact of the above-mentioned factors on student achievement, and to what extent do these beliefs diverge from findings of empirical research (i.e., the effect sizes of Hattie's research synthesis)?
b) What are the differences in the beliefs of novice and expert teachers on a latent level of meaningful factors of school effectiveness?

3. Methods

3.1 Sample

The sample comprised N = 729 participants (64% female); n = 358 were in-service ("expert") teachers and n = 371 pre-service ("novice") teachers. Teachers of the first group were in service at different schools in the federal states Schleswig-Holstein and Hamburg, Germany. Of these participants 53 % were women, the medium age was M = 52.3 years (SD = 8.8) ranging from 28 to 64 years. This subgroup included teachers from different types of schools (primary and secondary).The data from this subsample was collected in the context of professional development lectures for in-service teachers. Though attendance was not mandatory, the lectures were open for all teachers from the two states. The attending teachers can be considered true experts as many of them were in leadership positions at their schools (training supervision, school administration, etc.).

The pre-service teachers were university students enrolled in the first year of a Master of Education (M.Ed.) at a university in the northern part of Germany. To illustrate the background of our sample we briefly outline a typical teacher training program in Germany: At most German universities teacher training is composed of a three-year bachelor (B.A./ B.Sc.) program and a two-year master (M.Ed.) program. It includes the academic study of two scientific disciplines and didactics for the corresponding school subjects. In addition, students take a variety of courses in educational sciences. After university studies, students transfer to the more practical part of teacher training. They train teaching in schools for one to two years before they become proper teachers. Our pre-service teachers had completed a bachelor program that prepared them for graduate studies in teacher education. They had done practical training in schools for a period of six weeks in total; however, the bachelor program was clearly focused on the theoretical study of the two scientific disciplines. Only a small proportion of the degree was dedicated to introductory classes on educational sciences and didactics. Thus, we can assume that their prior knowledge concerning these subjects was not very advanced. The percentage of female participants in this group was 71%. Their mean age was M = 24.7 years (SD = 2.5), ranging from 23 to 39 years. The data was collected during a lecture on psychology in education, in which all students enrolled in the M.Ed. were required to participate.

3.2 Procedure

In order to assess teachers' beliefs about the effectiveness of factors for students' achievement a questionnaire was developed based on 16 determinants of student learning (see Table 2) selected from Hattie (2009). Criteria for the selection of items were the coverage of a large range of effect sizes (d = .01-.73) and

the coverage of the a priori categories *school*, *teaching* and *student* from Hattie's study. These categories were chosen as a focus of our study since there seemed to be the highest consensus about the extent of their impact on student learning among school effectiveness studies. Moreover, we selected those variables that we assumed even inexperienced university students would be familiar with, as most of them are also a frequent issue in political and academic discourse.

Table 1

Intervals of effect sizes and their interpretations by Hattie (2009) and Köller (2012)

Range of effect sizes	Interpretation by Hattie (2009)	Interpretation by Köller (2012)
$d < 0$	reverse effects	harmful
$0 \leq d < 0.15/0.2$	developmental effects	not harmful, not helpful
$0.15/0.2 \leq d < 0.4$	teacher effects	a little helpful
$0.4 \leq d < 0.6$	zone of desired effects	helpful
$d \geq 0.6$		very helpful

The questionnaire was administered in the context of a lecture on Hattie's study. First of all, participants were introduced to the concept of meta-analysis in general and the design of Hattie's research synthesis in particular. Subsequently, they were familiarized with the concept of Cohen's *d* and its practical implications: The formula $d = (M_{test} - M_{control}) / SD_{pooled}$ was presented to the teachers and explained in detail with the help of examples. In order to give practical significance to the rather abstract notion of effect size the interpretation of effect sizes as presented in the rightmost column of Table 1 was introduced and displayed during the completion of the questionnaire.

The participants were asked to estimate the impact of each of the factors on a scale of effect sizes ranging from $d = -0.4$ to $d = 1.0$. The precise instruction was: "Please estimate the effect of each of the factors below on students' achievement", followed by the list of variables. Participants were briefly familiarized with the 16 factors, that is, short comments were given on what was meant by the factors.

3.3 Statistical analyses

First, group means were calculated on the level of the 16 manifest variables and further analyzed by comparing them to the effect sizes of Hattie (2009). This was achieved by calculating the correlation coefficient Pearson's *r* for each person's rating profile with the distribution of Hattie's effect sizes. The coefficients were transformed by Fisher's *z* in order to approximate constant variance for all values. The use of Fisher's *z* transformation is recommended when averaging correlation coefficients as the distribution of *r* is skewed (Silver & Dunlap, 1987). The resulting Fisher's *z* coefficients were aggregated per group (pre- and in-service teachers) and the resulting means (M_z) of the two groups were compared using an independent-sample t-test. Thus, we could determine the difference in pre- and in-service teachers in terms of congruence with Hattie's results.

Second, we examined whether ratings on several individual items could be aggregated on a higher level, that is, in a latent variable model. Confirmatory factor analysis (CFA) was carried out in M*plus* 7 (Muthén & Muthén, 1998-2012) in order to analyze the underlying latent structure of the data, which then served as a basis for comparisons between pre- and in-service teachers on the reduced number of meaningful categories on a higher level. The assumed factor structure was based on Hattie's (2009) a priori categorization of the variables. The factor *teaching* contained variables from Hattie's categories *teaching* and *teacher*, the factor *school* contained variables from Hattie's category *school*, and the factor *student* contained variables from Hattie's category *student*. Modification indices indicated, however, that the factor *teaching*

was to be split in two (*teaching* and *achievement*). In the final measurement model we specified four latent variables using maximum likelihood estimation (ML). The latent variables were allowed to correlate and some error terms of manifest variables were allowed to covary if considered plausible. Missing data (< 5%) were estimated with the help of the full information maximum likelihood (FIML) procedure.

Subsequently, in order to allow for meaningful comparisons between the subgroups, measurement invariance was tested using a multiple-group modeling approach (Meredith & Teresi, 2006). For a comparison of latent means across the two groups (pre- vs. in-service teachers) at least partial scalar invariance is required (Byrne, Shavelson & Muthén, 1989). In multiple group analysis, when the specified model includes a mean structure, both the intercepts and factor loadings of the continuous factor indicators are held equal across groups to specify (scalar) measurement invariance. The intercepts of the factors are fixed at zero in the first group and are free to be estimated in the other groups. Thus, differences between the two groups can be determined based on the latent factors.

4. Results

4.1 Descriptive statistics and item level analysis

Table 2 shows descriptive statistics for the two groups – pre- and in-service teachers – on item level as well as Hattie's (2009) research results. The factors are ranked by the size of their effect (*d*) on student achievement as found by Hattie in his metasynthesis. In the following we elaborate on the descriptive results, especially on those variables with the highest and lowest ratings in each group. Both groups seemed to believe in the importance of student variables, as they both showed the highest means on the factors *motivation* and *attitude*. Ranked third by both pre- and in-service teachers was *feedback*. *Multi-grade/age learning* (in-service group) and *direct instruction* (pre-service group), respectively, had the lowest ratings of all 16 factors. Pre-service teachers had significantly higher effect sizes for the variables *feedback, prior achievement, motivation, attitude, class size, co-/team teaching, within-class grouping*, and *open learning*. In-service teachers' beliefs showed higher effect sizes for *direct instruction*, *high expectations*, and *self-concept*.

Table 2

Hattie's effect sizes (d), group means (M_{group}), and standard deviations (SD)

Factors	d_{Hattie}	$M_{in\text{-}service}$	$M_{pre\text{-}service}$
Feedback	.73	.55 (.23)	.62 (.24)[a]
Teaching meta-cognitive strategies	.69	.53 (.25)	.53 (.26)
Prior achievement	.67	.32 (.23)	.39 (.24)[a]
Professional development	.62	.40 (.23)	.42 (.25)
Direct instruction	.59	.28 (.23)[a]	.18 (.25)
Motivation	.48	.63 (.24)	.73 (.22)[a]
Expectations	.43	.36 (.25)[a]	.23 (.29)
Self-concept	.43	.55 (.21)[a]	.48 (.24)
Attitude	.36	.56 (.24)	.66 (.21)[a]
Frequent/effects of testing	.34	.34 (.24)	.34 (.26)

Class size	.21	.34 (.31)	.59 (.30)[a]
Co-/team teaching	.19	.37 (.27)	.45 (.28)[a]
Within-class grouping	.16	.48 (.27)	.61 (.24)[a]
Problem-based learning	.15	.52 (.23)	.52 (.25)
Multi-grade/age classes	.04	.19 (.26)	.22 (.27)
Open learning	.01	.29 (.27)	.37 (.27)[a]

[a] superscript characters indicate statistically significant ($p < .01$) higher mean effect sizes for the respective group (results of two-sample independent t-tests using Bonferroni correction to account for the multiple comparisons problem)

We determined bivariate correlations for each teacher's rating profile on the one hand and Hattie's results on the other by calculating Pearson's *r* for each person. The coefficients were transformed by Fisher's *z* and aggregated per group. These means were then compared using an independent-sample *t*-test. For the pre-service teachers the mean (Fisher *z*-transformed) correlation with Hattie's *d* was M_z=.06 (*SD* =.32), for the in-service teachers it was M_z = .23 (*SD* = .35). These group differences were statistically significant ($t[715] = 7.12$; $p < .001$; $d = .51$), indicating a substantially higher degree of conformity of the experts' ratings with Hattie's results.

4.2 CFA and multiple-group analysis

A priori, for the item pool selected for this study we assumed three latent factors based on Hattie's categorization of indicators: *school*, *teaching/teacher* and *student* ($\chi^2[95]$=586.36; CFI=.85; RMSEA=.08; TLI=.81; SRMR=.08). Empirically, however, a four-dimensional structure resulted in a better model fit ($\chi^2[92] = 340.85$; CFI = .92; RMSEA = .06; TLI = .90; SRMR = .05). The factor *teaching* was split in two, separating the strongly achievement-focused indicators from more specific instructional teaching behaviors. The improvement in goodness-of-fit indices was substantial (ΔCFI > .01; ΔRMSEA > .015) (Cheung & Rensvold, 2002), so we decided on the four-dimensional model (see Table 3). Residual correlations were allowed for some indicators with substantial covariance that was not explained by the latent factor.

All items loaded significantly ($p < .001$) and almost all items loaded substantially ($\lambda \geq .4$) on one of the latent factors. The only items with factor loadings slightly below the minimum value were prior achievement ($\lambda = .37$) and multi-grade/age classes ($\lambda = .37$). The former is the only indicator for the factor student that focuses on cognitive rather than motivational aspects of a student's academic prerequisites. This might explain the low factor loading. The latter may have been a difficult concept for many of the participants as by far not all teachers encounter this instructional challenge throughout their careers.

Due to high modification indices, residual correlations were allowed for six item pairs (*multi-grade/age classes* with *open learning* and *co-/team teaching*, *class size* with *co-/team teaching* and *within-class grouping*, *teaching meta-cognitive strategies* with *feedback*, *motivation* with *attitude*). The majority of these modifications were performed within the factor *structure*. They seemed to be theoretically sound as certain infrastructural conditions of schooling (*class size*, *multi-grade/age classes*) are strongly associated with or even demand certain surface-structural aspects of learning (*open learning*; *co-/team teaching*; *within-class grouping*). *Teaching meta-cognitive strategies* and *feedback* are both direct and concrete instructional measures of the teacher, *motivation* and *attitude towards subject* refer to very similar student-internal constructs (as opposed to the other respective indicators). The allowed covariances were the same for both models (three and four latent factors) that were tested.

Table 3

Standardized factor loadings matrix of the CFA

		Teaching	Achievement	School	Student
1	Feedback	0.62	-	-	-
2	Teaching meta-cognitive strategies	0.68	-	-	-
3	Professional development	0.70	-	-	-
4	Problem-based learning	0.69	-	-	-
5	Direct instruction	-	0.52	-	-
6	Expectations	-	0.69	-	-
7	Frequent/effects of testing	-	0.48	-	-
8	Multi-grade/age classes	-	-	0.37	-
9	Open learning	-	-	0.65	-
10	Class size	-	-	0.51	-
11	Co-/team teaching	-	-	0.56	-
12	Within-class grouping	-	-	0.77	-
13	Motivation	-	-	-	0.67
14	Prior achievement	-	-	-	0.37
15	Self-concept	-	-	-	0.50
16	Attitude towards subject	-	-	-	0.57

In the following we will explain the four dimensions in more detail: The factor *school* comprises infrastructural conditions of schooling (*class size; multi-grade/age classes*) and the surface-structure of learning (*open learning; co-/team teaching; within-class grouping*). The factor *teaching* contains manifest variables concerning instructional methods (*feedback; teaching meta-cognitive strategies; problem-based learning*) and the teacher (*professional development*), while the factor *achievement* emphasizes achievement-focused and teacher-centered variables (*direct instruction*; *teacher expectations; frequent/effects of testing*). Student-internal prerequisites (*motivation*; *self-concept; prior achievement; attitude*) constitute the factor *student*.

Table 4

Correlation matrix of latent factor model

	Achievement	School	Student
Teaching	.41**	.70**	.66**
Achievement		-.16**	.17*
School			.78**

**p < .01; *p < .05

Table 4 shows bivariate correlations of the four latent variables, which were all statistically significant. All coefficients were positive apart from the one between achievement and school, which showed a negative relationship. The intercorrelations were strongest for teaching/school, teaching/student, and school/student. The factor achievement consistently showed the weakest relationships with all the other

factors. These results indicate that, in general, teachers had a tendency to believe in either high or low effect sizes for factors of school effectiveness. However, achievement-related variables seemed to be the exception.

Table 5

Measurement invariance across groups (pre- and in-service teachers)

Model	Parameters constrained	χ^2	*df*	*CFI*	*TLI*	*RMSEA*	*SRMR*
1	None (configural invariance)	405.22	184	.932	.912	.058	.055
2	FL (metric invariance)	428.09	196	.929	.913	.057	.057
3	FL, IL (scalar invariance)	627.50	208	.872	.852	.075	.069
3b	FL, IL (partial scalar invariance)	448.51	205	.925	.913	.058	.058

Note. FL = factor loadings. II = item intercepts. For model identification in model 1 and 2 (item intercepts freely estimated) latent means were fixed to zero. *CFI* = Comparative Fit Index. *TLI* = Tucker-Lewis index. *RMSEA* = root mean square error of approximation. *SRMR* = standardized root mean square residual.

The model showed partial scalar invariance (strong measurement invariance; see Table 5), which indicated that factor structure and factor loadings as well as item intercepts were equal for both groups. Pre- and in-service teachers attributed the same meaning to the latent constructs and the levels of the underlying items. Thus, the proposed factor model can be assumed to represent the belief structure of pre-service as well as in-service teachers. Strong measurement invariance allows for the comparison of latent group means. In the latent mean structure analysis the pre-service teachers were chosen as a reference group so that the difference in means between pre- and in-service teachers on each construct equals the mean of the non-reference group (in-service teachers). The means of the in-service teachers are as shown in Table 6. They valued the achievement-related factor considerably higher, while rating the effects of infra-/surface-structure and student-internal variables lower than pre-service teachers. The group differences concerning the factor teacher/teaching were not significant.

Table 6

*Mean group differences in latent variables**

Factor	$M_{\text{difference}}$	*p*
Teaching	-0.12	*ns*
Achievement	0.67	<.001
School	-0.56	<.001
Student	-0.70	<.001

*$M_{\text{pre-service}} = 0$; $M_{\text{in-service}} = M_{\text{difference}}$

5. Discussion

5.1 Summary

The present paper deals with expert and novice teachers' beliefs about school effectiveness. We investigated the differences between (a) teachers' beliefs versus findings of school effectiveness research (cf. Hattie, 2009), and (b) expert versus novice teachers' beliefs. For this purpose pre- and in-service teachers

were asked to rate the effect sizes of several determinants of student achievement. Profile correlations were aggregated and compared in terms of similarity to recent empirical findings (Hattie, 2009). We found significant differences between pre- and in-service teachers, as the latter showed a stronger overall congruence with Hattie's results. Subsequently, data were combined by using a four-dimensional CFA-model with the latent factors *school*, *teaching*, *achievement*, and *student*. Partial measurement invariance could be established allowing the comparison of latent factor means of the two groups. In-service teachers showed higher means in achievement-focused variables (e.g., *direct instruction*, and *high expectations*) and lower means in variables concerning the infra- and surface-structural conditions of schooling (e.g., *class size, co-/team teaching, within-class grouping*, and *open learning*) as well as student-internal variables (e.g., *prior achievement*, *motivation*, and *attitude*).

The structure of teachers' beliefs concerning school effectiveness seemed to resemble the a priori categorization of relevant research studies. The assumed categorical structure (*school*, *teaching*, and *student*) used by Hattie (2009) to organize his metasynthesis was largely met by the data. An exception was the separation of the factor teaching into teaching and achievement. Teachers seemed to distinguish two types of instructional preferences: one that foregrounds the support of students' learning (*feedback*, *teaching meta-cognitive strategies*) and one that focuses mainly on cognitive achievement (*high expectations*, *frequent/effects of testing*). While school effectiveness research has shown that both have a strong positive influence on student achievement (see chapter 1.3), teachers undervalued instructional choices that point directly and explicitly at academic achievement.

One key finding of this study was that in-service experience and training seem to be associated with teachers' beliefs about school effectiveness. The beliefs of experienced teachers were more consistent with empirical results than those of novice teachers. This is in line with the current theoretical expert(-novice) paradigm, according to which teachers develop expertise in the course of their education and career. First of all, the ratings of our in-service teachers suggested that they value a kind of activating, teacher-directed instruction, which is supposed to affect the deep structure of classroom learning more than the pre-service teachers do. Second, in comparison to the novices they believed infra- and surface-structural variables to be not as relevant. This hierarchy is also emphasized by the results of school effectiveness research as outlined in chapter 1.3. This compliance indicates that expert teachers are more competent in assessing the effects of a range of variables than novice teachers. Furthermore, it shows that educational research appears not as far from a teacher's perceived everyday reality as is often suggested.

In turn, we must acknowledge that some of the beliefs concerning influences on student achievement – particularly (but not only) those of pre-service teachers – diverge from empirical findings quite dramatically. Similarly to Weinstein (1989), we found that affective (e.g. *motivation*, *attitude*) and social (e.g. *within-class grouping*) variables were overestimated, while cognitive variables (e.g. *direct instruction*, *prior achievement*) were underestimated, especially by the pre-service teachers, but also – to a lesser extent – by in-service teachers. This insight is quite valuable considering the impact of beliefs on classroom instruction. Hence, our results call for a paradigmatic change in the way teachers are trained. In the following we consider the limitations of this study before we conclude with its strengths and practical implications for the field of teacher education.

5.2 Limitations

One problem with the interpretation of our results was the epistemological status of the information given by the participants: Did we actually elicit their beliefs or rather their theoretical knowledge about what should be correct. As we mentioned above, the confrontation with objective data, such as the findings of empirical investigations, does not necessarily lead to a change in individual beliefs: *Knowing* something does not equal *believing* in it. So the response of the teachers to a certain item may not have shown their beliefs but instead, represent the knowledge they have about school effectiveness research. Hence, the epistemological distinction between beliefs and knowledge that we addressed above could not be followed

through completely. Similarly, the issues of tacit knowledge, its accessibility and the reciprocal relationship of implicit and explicit knowledge are highly relevant topics that go beyond the scope of this study.

Even if we assume that we are dealing with teachers' beliefs (not knowledge), we cannot be certain that those beliefs are actually put into practice. In our report of prior research on teachers' beliefs we pointed out that beliefs are considered to affect classroom processes and, in turn, student outcomes. However, the assumption that teachers apply everything they believe to classroom practice would go too far (belief-behavior gap; cf. Sheeran, 2002). So a general identification of "good beliefs" with "good teacher" is too simplistic. This issue requires further research that examines the relevance of teachers' beliefs (concerning factors of school effectiveness) for instructional choices and student achievement.

Another drawback is that in the present study Hattie's findings served as a kind of "quasi-reality". Despite all its merits, as a synthesis of meta-analyses it also poses some methodological difficulties (for a more extensive discussion see Terhart, 2011), which limits the interpretability of the discrepancy between "real" and teacher-estimated effect sizes. Thus, even though we concentrated on those variables that have shown similar results in other school effectiveness studies, the comparison with Hattie's results is to be interpreted with caution. The same holds for the comparison of pre- and in-service teachers: As our data was not longitudinal, strictly speaking we cannot interpret the discrepancy between the groups as a development or acquisition of competence. Additionally, in order to validate and stabilize the latent factor structure further investigations with a larger number of variables from Hattie's study are needed.

The generalizability of our results to other countries is limited for two different reasons: Firstly, the intercultural generalizability of Hattie's study is already questionable. It mainly relies on research findings from Anglophone countries, not all of which are equally relevant for education in Germany and for the beliefs of German teachers. Secondly, we also have to take into account that our sampling was restricted to the German education system and, moreover, to one specific teacher training program. Beliefs are rather likely to differ in terms of structure- and content-related aspects of teacher education in different countries and institutions. Furthermore, they are subject to the more general cultural and academic situation in a certain time and place. For example, the acceptance and appreciation of empirical research may differ considerably from country to country and from faculty to faculty. In case of this study, the fact that German teacher training programs generally focus on the study of two scientific disciplines rather than on educational science and field experience might impact the beliefs of teachers. The question of differences in teachers' beliefs according to differences in their (culture- and program-specific) education and practice would be an interesting topic for future research.

Our research was restricted to teachers' beliefs concerning cognitive achievement. Unfortunately, Hattie's work does not identify determinants of motivation, attitude, self-concept or other affective variables. A study that examines determinants of affective outcomes to the extent that Hattie did this for cognitive outcomes is still missing in educational effectiveness research. Thus, there is no basis for research on teachers' beliefs concerning those variables yet.

5.3 Strengths and educational implications

Especially the beliefs of pre-service teachers differed significantly from the results of Hattie's research synthesis. Of course neither Hattie's findings nor the beliefs of expert teachers can be taken as ultimately true or as factual reality. However, they both emphasize similar aspects of schooling: the role of the teacher as an *activator* (rather than a *facilitator*), the importance of academic achievement and the comparably little significance of structural conditions. If teacher educators address these issues explicitly and confront their students with their own beliefs as well as with the findings of school effectiveness research, they can help (prospective) teachers to focus on what has yet been shown to work best in school.

The fact that pre-service teachers' beliefs diverged from empirical findings more strongly than those of experienced teachers suggests that learning opportunities in the field do make a difference. The experience

that teachers gain in their years of classroom practice seems to affect their judgment in a way that is beneficial for their belief systems. One could argue that the constant feedback they get from their students' performance in terms of (successful and unsuccessful) interventions helps them challenge and adjust their beliefs where necessary. In-service teachers' focused on instructional strategies and factors that support the deep structure of learning might thus be a reaction to their (more or less systematic) observation and monitoring of what actually works in their own classroom. Pre-service teachers, however, are missing this direct feedback in terms of actual student outcomes as their confrontation with actual classroom situations is very limited. Holding on to established beliefs might be a result of them not being challenged by reality in the field. Moreover, one could expect novice teachers to be quite overwhelmed by their first tentative efforts in teaching. The demands they have to meet in the classroom are manifold and they need to concentrate on various things at once. In such a state, focusing on the surface structure of learning seems easier than focusing on deep-structural aspects. Only with substantial practice and experience, when other processes come to them more naturally and intuitively, teachers get the opportunity to pay attention to those instructional details that have actually been shown to work in school.

Early and regular work experience in school during teacher training might aid the acquisition of the necessary professionalism, as it presents the opportunity to familiarize pre-service teachers with real classroom situations and their role as a teacher. However, this should be realized only with respect to the current state of research as we have little reason to assume that practical school experiences for pre-service teachers automatically lead to better teaching abilities or a better understanding for the purposes and consequences of teaching. (Tabachnik et al., 1979-1980). Practical experience in school does not automatically make better teachers. This also applies to their beliefs: Studies on short-term work experiences during teacher training has shown the resilience of teachers' beliefs to change (Hascher, 2012; Richardson, 1996). In order to avoid this misguided process, these early experiences need to be instructed and accompanied by professional teacher trainers. If planned and exercised carefully, practical teaching experience during teacher training at university can lay a solid foundation for a teacher's career.

The other central finding of this study was the fundamental discrepancy of teachers' beliefs and empirical evidence from school effectiveness research. To some extent, this might be due to shortcomings in teacher training programs to convey to future teachers the importance of evidence-based practice. The Australian educational scholar and administrator Michele Bruniges puts the lack of data usage in the teaching profession into the following words:

> "A Greek philosopher might suggest that evidence is what is observed, rational and logical; a Fundamentalist – what you know is true; a Post Modernist – what you experience; a Lawyer – material which tends to prove or disprove the existence of a fact and that is admissible in court; a Clinical Scientist – information obtained from observations and/or experiments; and a teacher – what they see and hear" (Bruniges, 2005; p. 102).

While systematic observation and monitoring of students' learning processes are very desirable actions to be taken by teachers, "seeing and hearing" should not be the only sources for their professional choices and actions. Our study supports the claim that teachers rarely rely on available research evidence. Their assessment of what actually works in school rather seems to be guided by subjective experiences that are usually gained in the isolation of their own classrooms.

But it would certainly be wrong to lay all the blame on the teachers: What we need is an evidence-based culture of improvement in teaching and learning. In order to achieve this goal, three professions in the field of education need to assume responsibility: researchers, teacher trainers, and teachers themselves. First of all, educational researchers are confronted with the issue of making their findings available to teachers. More often than it is already done, they should break rather abstract studies down to what is of practical relevance for the field. Such efforts may counteract aversion to empirical research on the side of the teachers. With his follow-up book "Visible Learning for Teachers", Hattie sets a good example for this kind of transfer. Secondly, those who educate and train pre-service teachers need to make sure their students are

familiar with relevant research findings, can interpret them appropriately, and have the necessary skills to implement them in school. In addition to assessing students' knowledge, they should be attentive to their beliefs and make room for critical discussion of empirical versus anecdotal evidence. Explicitly addressing the issue of teachers' beliefs and confronting (future) teachers with cognitive dissonance might support a critical reflection and examination of existing beliefs. Last but not least, it is a necessity for teachers themselves to stay in touch with research communities in order to understand current developments and to constantly reflect on their beliefs in comparison with crucial evidence provided by researchers.

Keypoints

- Teachers' beliefs diverge from empirical evidence
- Expert teachers' beliefs diverge from novice teachers' beliefs
- Expert teachers show more congruence with empirical evidence than novice teachers
- Expert teachers believe in the effectiveness of achievement-related variables
- Novice teachers believe in the effectiveness of structural and student factors

References

Allen, L. Q. (2002). Teachers' pedagogical beliefs and the standards for foreign language learning. *Foreign Language Annals, 35*, 518-529. http://dx.doi.org/10.1111/j.1944-9720.2002.tb02720.x

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Barcelos, A. M. F. (2003). Researching beliefs about SLA: A critical review. In P. Kalaja and A. M. F. Barcelos (Eds.), *Beliefs about SLA: New research approaches* (pp. 7-33). Dordrecht: Kluwer Academic Publishers. http://dx.doi.org/10.1007/978-1-4020-4751-0_1

Bates, A. B., Latham, N. & Kim, J. (2011). Linking preservice teachers' mathematics self-efficacy and mathematics teaching efficacy to their mathematical performance. *School Science and Mathematics*. *111*, 325-333. http://dx.doi.org/10.1111/j.1949-8594.2011.00095.x

Bauer, J. & Prenzel, M. (2012). European teacher training reforms. *Science*, *336*, 1642-1643. http://dx.doi.org/10.1126/science.1218387

Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft, 9*, 469-520. http://dx.doi.org/10.1007/s11618-006-0165-2

Berliner, D. (2004). Describing the behavior and documenting the accomplishments of expert teachers. *Bulletin of Science, Technology and Society*, *24*, 200-214. http://dx.doi.org/10.1177/0270467604265535

Borg, M. (2001). Teachers' Beliefs. *ELT Journal, 55*, 186-188. http://dx.doi.org/10.1093/elt/55.2.186

Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching, 36*, 81-109. http://dx.doi.org/10.1017/S0261444803001903

Bromme, R. (1997). Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers. In F. E. Weinert (Ed.), *Psychologie des Unterrichts und der Schule. Enzyklopaedie der Psychologie* (Vol. 3, pp. 177-212). Goettingen: Hogrefe.

Bromme, R. (2003). On the limitations of the theory metaphor for the study of teachers' expert knowledge. In M. Kompf & P. Denicolo (Eds.). *Teacher thinking twenty years on: Revisiting persisting problems and advances in education* (pp. 283-294). Liss, Nl: Swets & Zeitlinger.

Bruniges, M. (2005). An evidence-based approach to teaching and learning. http://research.acer.edu.au/research_conference_2005/15.

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, *105*, 456-466. http://dx.doi.org/10.1037/0033-2909.105.3.456

Calderhead, J. (1996). Teachers: Beliefs and knowledge. In D. Berliner & R. Calfee (Eds.), *Handbook of Educational Psychology* (pp. 709-725). New York: Macmillan.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing MI. *Structural Equation Modeling*, *9*, 235-55. http://dx.doi.org/10.1207/S15328007SEM0902_5

Clark, C. M. (1988). Asking the right questions about teacher preparation: Contributions of research on teacher thinking. *Educational Researcher, 17*, 5-12. http://dx.doi.org/10.3102/0013189X017002005

Cochran-Smith, M., & Zeichner, K. (Eds.). (2005). *Studying teacher education: The report of the AERA panel on research and teacher education*. Mahwah: Lawrence Erlbaum Associates.

Coleman, J. S., Campbell, E. Q., Hobson, C. F., McPartland, J., Mood, A. M., Weifeld, F. D., et al. (1966). *Equality of educational opportunity*. Washington, D.C.: U.S. Government Printing Office.

Cotton, K. (1995). *Effective schooling practices: A research synthesis*. 1995 Update. School Improvement Research Series. Northwest Regional Educational Laboratory.

Darling-Hammond, L., & Bransford, J. (Eds.). (2005). Preparing teachers for a changing world. What teachers should learn and be able to do. San Francisco: Jossey-Bass.

Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, P. Feltovich, and R. R. Hoffman, R. R. (Eds.). *Cambridge Handbook of Expertise and Expert Performance* (pp. 685-706). Cambridge, UK: Cambridge University Press.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptations to task constraints. *Annual Review of Psychology, 47*, 273-305. http://dx.doi.org/10.1146/annurev.psych.47.1.273

Ericsson, K. A., Charness, N., Feltovich, P. J., & Hoffman, R. R. (2006). The Cambridge handbook of expertise and expert performance. Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511816796

Fang, Z. (1996). A review of research on teacher beliefs and practices. *Educational Research*, 38, 47-65. http://dx.doi.org/10.1080/0013188960380104

Fraser, B. J., Walberg, H. J., Welch, W. W., & Hattie, J. A. (1987). Syntheses of educational productivity research. *International Journal of Educational Research, 11*, 147-252. http://dx.doi.org/10.1016/0883-0355(87)90035-8

Freeman, D. (2002). The hidden side of the work: Teacher knowledge and learning to teach. *Language Teaching, 35*, 1-13. http://dx.doi.org/10.1017/S0261444801001720

Haney, J. J., Lumpe, A. T., & Czerniak, C. M. (2003). Constructivist beliefs about the science classroom learning environment: Perspectives from teachers, administrators, parents, community members, and students. *School Science and Mathematics, 103*, 366-377. http://dx.doi.org/10.1111/j.1949-8594.2003.tb18122.x

Hart, L. C. (2004). Beliefs and perspectives of first-year, alternative preparation, elementary teachers in urban classrooms. *School Science & Mathematics*, *104*, 79-88. http://dx.doi.org/10.1111/j.1949-8594.2004.tb17985.x

Harvey, O. J. (1986). Belief systems and attitudes toward the death penalty and other punishments. *Journal of Psychology*, *54*, 143-159. http://dx.doi.org/10.1111/j.1467-6494.1986.tb00418.x

Hascher, T. (2012). Lernfeld Praktikum – Evidenzbasierte Entwicklungen in der Lehrer/innenbildung. [Learning setting student teaching – Evidence-based developments in teacher education]. *Zeitschrift für Bildungsforschung*, *2*, http://dx.doi.org/10.1007/s35834-012-0032-6

Hattie, J. (2003). *Teachers make a difference: What is the research evidence?* Paper presented at the Australian Council for Educational Research annual conference on building teacher quality.

Hattie, J. (2009). Visible learning. A synthesis of over 800 meta-analyses relating to achievement. London & New York: Routledge.

Hattie, J. (2012). *Visible learning for teachers*. London: Routledge.

Hogan, T., Rabinowitz, M., & Craven, J. A. (2003). Representation in teaching: Inferences from research of expert and novice teachers. Educational Psychologist, *38*, 235-247. http://dx.doi.org/10.1207/S15326985EP3804_3

Jencks, C., Smith, M. S., Ackland, H., Bane, M. J., Cohen, D., Grintlis, H., et al. (1973). *Inequality: A reassessment of the effect of family and schooling in America*. New York: Basic Books.

Köller, O. (2012). What works best in school? Hatties Befunde zu Effekten von Schul- und Unterrichtsvariablen auf Schulleistungen. [What works best in school? Hattie's findings concerning school and instructional effectiveness on student achievement]. *Psychologie in Erziehung und Unterricht, 59*, 72-78. http://dx.doi.org/10.2378/peu2012.art06d

Kunter, M., & Pohlmann, B. (2009). Lehrer. In J. Möller & E. Wild (Eds.), *Einführung in die Pädagogische Psychologie* (pp. 261-282). Berlin: Springer. http://dx.doi.org/10.1007/978-3-540-88573-3_11

Leatham, K. (2006). Viewing mathematics teachers' beliefs as sensible systems. *Journal of Mathematics Teacher Education, 9*, 91-102. http://dx.doi.org/10.1007/s10857-006-9006-8

Levine, D.K., & Lezotte, L.W. (1990*). Unusually effective schools: A review and analysis of research and practice*. Madison, Wise: Nat. Centre for Effective Schools Research and Development.

Lortie, D. (1975). *Schoolteacher: A sociological study*. Chicago: University of Chicago Press.

Meredith, W., & Teresi, J. A. (2006). An essay on measurement and factorial invariance. *Medical Care, 44*, 69-77. http://dx.doi.org/10.1097/01.mlr.0000245438.73837.89

Muijs, R. D., & Rejnolds, D. (2001). Teachers' beliefs and behaviors: What really matters. *Journal of Classroom Interaction*, *37*, 3-15.

Muthén, L.K. and Muthén, B.O. (1998-2012). Mplus user's guide. Seventh edition.

Los Angeles, CA: Muthén & Muthén.

National Board for Professional Teaching Standards (2002). *What teachers should know and be able to do*. Arlington.

Nespor, J. (1987). The role of beliefs in the practice of teaching. *Journal of Curriculum Studies, 19*, 317-328. http://dx.doi.org/10.1080/0022027870190403

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research, 62*, 307-332. http://dx.doi.org/10.3102/00346543062003307

Peterson, P., Fennema, E., Carpenter, T. P., & Loef, M. (1989). Teachers' pedagogical content beliefs in mathematics. *Cognition and Instruction, 6*, 1-40. http://dx.doi.org/10.1207/s1532690xci0601_1

Purkey, S.C., & Smith, M.S. (1983). Effective schools: A review. *The Elementary School Journal*, 83, 427-452. http://dx.doi.org/10.1086/461325

Richardson, V. (1996). The role of attitudes and beliefs in learning to teach. In J. Sikula, T. Buttery, & E. Guyton (Eds.), *Handbook of research on teacher education* (pp. 137-147). New York: Macmillan.

Ross, J. A. (1992). Teacher efficacy and the effect of coaching on student achievement. *Canadian Journal of Education, 17*, 51-65. http://dx.doi.org/10.2307/1495395

Ross, J. A. (1998). The antecedents and consequences of teacher efficacy. In J. Brophy (Ed.), *Advances in research on teaching, Vol. 7* (pp. 49-74). Greenwich, CT: JAI Press.

Sammons, P., Hillman, J., & Mortimore, P. (1995). *Key characteristics of effective schools: A review of school effectiveness research*. London: OFSTED.

Scheerens, J. (1992). *Effective schooling, research, theory and practice.* London: Cassell.

Scheerens, J. (2004). Review of school and instructional effectiveness research. Paper commissioned for the *EFA Global Monitoring Report 2005, The Quality Imperative.* UNESCO, 2005/ED/EFA/MRT/PI/44.

Scheerens, J., & Bosker, R. J. (1997). *The foundations of educational effectiveness*. Oxford: Elsevier Science Ltd.

Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research, 77*, 454-499. http://dx.doi.org/10.3102/0034654307310317

Sheeran, P. (2002). Intention-behavior relations: A conceptual and empirical review. In W. Strobe and M. Hewstone (Eds.) *European Review of Social Psychology*, Vol. 12. Chichester: Wiley, 1-30. http://dx.doi.org/10.1080/14792772143000003

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*, 4-14. http://dx.doi.org/10.3102/0013189X015002004
Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*, 1-22.
Silver, N. C., & Dunlap, W. P. (1987). Averaging correlation coefficients: Should Fisher's *z* transformation be used? *Journal of Applied Psychology, 72*, 146-148. http://dx.doi.org/10.1037/0021-9010.72.1.146
Staub, F. C., & Stern, E. (2002). The nature of teachers' pedagogical content beliefs matters for students' achievement gains: Quasi-experimental evidence. *Journal of Educational Psychology, 94*, 344-355. http://dx.doi.org/10.1037/0022-0663.94.2.344
Stuart, C., & Thurlow, D. (2000). Making it their own: Pre-service teachers' experiences, beliefs, and classroom practices, *Journal of Teacher Education, 51*(2), 112-121. http://dx.doi.org/10.1177/002248710005100205
Tabachnik, B. R., Popkewitz, T., & Zeichner, K. (1979-1980). Teacher education and the professional perspectives of student teachers. *Interchange*, *80*, 12-29. http://dx.doi.org/10.1007/BF01810816
Terhart, E. (2011). Has John Hattie really found the holy grail of research on teaching? An extended review of Visible Learning. *Journal of Curriculum Studies, 43*(3), 425-438. http://dx.doi.org/10.1080/00220272.2011.576774
Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive concept. *Teaching and Teacher Education, 17*, 783-805. http://dx.doi.org/10.3102/00346543068002202
Tschannen-Moran, M., Woolfolk Hoy, A., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68*, 202-248.
Verloop, N., Van Driel, J., & Meijer, P. (2001). Teacher knowledge and the knowledge base of teaching. *International Journal of Education Research, 35*, 441-461. http://dx.doi.org/10.1016/S0883-0355(02)00003-4
Walberg, H. J. (1986). Synthesis of research on teaching. In M. C. Wittrock (Ed.), *Handbook of Research on Teaching*. New York: Macmillan.
Wang, M. C., Haertel, G. D. & Walberg, H. J. (1990). What influences learning? A content analysis of review literature. *Journal of Educational Research, 84*, 30-43. http://dx.doi.org/10.1080/00220671.1990.10885988
Wang, M. C., Haertel, G. D. & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research, 63*, 249-294. http://dx.doi.org/10.3102/00346543063003249
Weinstein, C. S. (1989). Teacher education students' preconceptions of teaching. *Journal of Teacher Education, 40*, 53-60. http://dx.doi.org/10.1177/002248718904000210
Wenden, A. L. (1999). An introduction to metacognitive knowledge and beliefs in language learning: Beyond the basics. *System, 27*, 435-441. http://dx.doi.org/10.1016/S0346-251X(99)00043-3
Woods, D. (1996). *Teacher cognition in language teaching: Beliefs, decision-making, and classroom practice*. Cambridge: Cambridge University Press.
Woolfolk, A. E., & Hoy,W. K. (1990). Prospective teachers' sense of efficacy and beliefs about control. *Journal of Educational Psychology*, *82*, 81-91. http://dx.doi.org/10.1037/0022-0663.82.1.81
Woolfolk, A. E., Rosoff, B., & Hoy, W. K. (1990). Teachers' sense of efficacy and their beliefs about managing students. *Teaching and Teacher Education, 6*, 137-148.
Woolfolk Hoy, A., & Davis, H. A. (2006). Teacher self-efficacy and its influence on the achievement of adolescents. In F. Pajares & T. Urdan (Eds.), *Self-efficacy of adolescents* (pp. 117-137). Greenwich, Connecticut: Information Age Publishing. http://dx.doi.org/10.1016/0742-051X(90)90031-Y
Woolfolk Hoy, A., Davis, H., & Pape, S. J. (2006). Teacher knowledge and beliefs. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of Educational Psychology* (2 ed., pp. 715-737). Mahwah, NJ: Erlbaum.
Yook, C. M. (2010). Korean teachers' beliefs about English language education and their impacts upon the Ministry of Education-initiated reforms. *Applied Linguistics and English as a Second Language Dissertations.* Paper 14.
Zell, E., & Krizan, Z. (2014). Do people have insight into their abilities? A metasynthesis. *Perspectives on Psychological Science, 9*, 111-125. http://dx.doi.org/10.1177/1745691613518075

Zheng, H. (2009). A review of research on pre-service teachers' beliefs and practices. *Journal of Cambridge Studies*, *4*, 73-81.