Fostering inquiry-based pedagogy in primary school:

A longitudinal study into the effects of a two-year school improvement project

Tim Post

Fostering inquiry-based pedagogy in primary school:
A longitudinal study into the effects of a two-year school
improvement project

Tim Post

FOSTERING INQUIRY-BASED PEDAGOGY IN PRIMARY SCHOOL: A LONGITUDINAL STUDY INTO THE EFFECTS OF A TWO-YEAR SCHOOL IMPROVEMENT PROJECT

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof. dr. T. T. M. Palstra,
volgens besluit van het College voor Promoties
in het openbaar te verdedigen
op woensdag 11 december 2019 om 16:45 uur

door

Tim Post

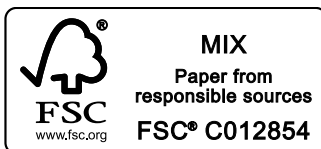geboren op 29 januari 1984
te Nieuwegein

Dit proefschrift is goedgekeurd door de promotor:
Prof. dr. J. H. Walma van der Molen

UNIVERSITY OF TWENTE.

TECH YOUR FUTURE
Centre of Expertise TechniekOnderwijs

ico

FSC
www.fsc.org
MIX
Paper from responsible sources
FSC® C012854

GRADUATION COMMITTEE

Chair: Prof. dr. T. A. J. Toonen

Supervisor: Prof. dr. J. H. Walma van der Molen

Members: Prof. dr. A. J. M. de Jong
                Prof. dr. P. C. Meijer
                Prof. dr. M. J. de Vries
                Prof. dr. E. T. Bohlmeijer
                Dr. H. W. Steenbeek

Table of contents

Chapter 1

1   Introduction

This dissertation describes the results of a five-year research project (0,84fte) that was funded by TechYourFuture, the Dutch center of expertise in technology education. TechYourFuture is a partnership between the University of Twente, Saxion, and Windesheim, which aims at encouraging young people in The Netherlands to opt for a technology-oriented study and career path. The project started in September 2013 in response to various national education policy changes in Dutch primary education.

Context of the study

International education policy documents increasingly promote the implementation of primary science and technology (S&T) school curricula (OECD, 2015; Osborne & Dillon, 2008). Scientific and technological innovations take place in a rapidly increasing rate and lead to the ongoing transformation of labor markets and societal structures (World Economic Forum, 2018). All young people will therefore have to become sufficiently familiar with S&T to be able to fully participate as future citizens and professionals in society (National Research Council, 2012). Furthermore, research indicates that children's natural interest for studying and working in S&T-related fields decreases if they have not developed affinity with S&T by the end of primary education (Turner & Ireson, 2010).

In the last two decades, the Dutch government has therefore promoted the widespread implementation S&T teaching in primary education. To this end, various national projects were funded, among most noticeably, Verbreding Techniek Basisonderwijs (VTB) (in English: Broadening Technology Education in Primary Education) and VTB-Pro (extended focus on teacher professionalization) by Platform Betá Techniek (in English: Platform Beta Technology).

The VTB project was initiated in the year 2004, which promoted the implementation S&T education in 2500 Dutch primary schools (about a third of the total number of primary schools in The Netherlands) by the end of the year 2010. Schools were provided financial support to appoint a S&T coordinator among their staffs and to purchase science lesson examples for teaching S&T. Although the VTB project generated significant interest among many primary schools, the project outcomes were limited. Results indicated that most primary teachers lacked the

competency to teach S&T using the provided lesson materials (De Vries, Van Keulen, Peters, & Walma van der Molen, 2011). S&T lessons often remained overly prescriptive and incidental, which left pupils little opportunity for authentic and substantial inquiry into the S&T domains. Research indicates that, in many Western countries, most primary teachers feel incapable of teaching S&T and thus often shy away from such teaching (Osborne & Dillon, 2008). Teacher professional development has proven effective in helping primary teachers acquire S&T teaching competency (Syer, Chichekian, Shore, & Aulls, 2012; Walan, Mc Ewen, & Gericke, 2016).

Therefore, in the year 2008, the VTB-Pro project was started to offer pre-service and in-service primary teachers (subsidized) teacher professionalization to develop S&T teaching competency (De Vries et al., 2011). It was further expected that, by having received the training, the participants would encourage members of their own school teams to develop similar competency over time, such as through teacher collaboration. Numerous other initiatives were taken to further support implementation, such as the option for teachers to receive assistance by professionals who worked or had worked in S&T-related industry (e.g., to assist them with teaching S&T) or to receive support with teaching S&T (e.g., extra lesson materials, best practices) through various regional support centers at universities in the country (in Dutch referred to as: 'wetenschapsknooppunten').

The VTB-Pro project proved successful in several important areas (De Vries et al., 2011): the teacher training helped participants feel more confident about teaching S&T and thereby encouraged them to more frequently teach S&T in their own classrooms. However, evaluations also indicated that relatively few primary teachers opted for participating in the training (about 10 percent of all primary teachers in The Netherlands), which suggest that many schools did not perceive the relevance and urgency of such training. Moreover, among the participants, few teachers appeared successful in involving other members of their school (including their school principals) with the school-wide implementation of S&T education. In turn, they received little organizational support and guidance from their school leadership to reform their practices, such as opportunities for classroom experimentation and teacher collaboration. The above approaches to the promotion of S&T education in Dutch primary schools thus left room for improvement.

In response to the above project outcomes, in the year 2012, the Council for Primary Education and Platform Bèta Techniek established a special Exploratory Committee comprised of various Dutch education experts. The committee was

charged with the task to recommend evidence-informed guidelines for an improved and sustained implementation of S&T teaching in Dutch primary education (for the full report on these guidelines, please see Verkenningscommissie, 2013). International research proposed various changes to education to successfully introduce S&T education in primary schools, especially in the domains of scientific literacy (e.g., Lumpe, Haney, & Czerniak, 2000), science education (e.g., Osborne & Dillon, 2008), creativity (e.g., Lucas, Claxton, & Spencer, 2013), and twenty-first century learning (e.g., Pellegrino & Hilton, 2012). In addition, school improvement research indicated various school factors that are important for successful school practice reform (e.g., Thoonen, Sleegers, Oort, Peetsma, & Geijsel, 2011).

Based on the above literatures, the Exploratory Committee proposed the following three main guidelines: (1) the adoption of a *broader* meaning of 'S&T teaching' in primary education, which would provide teachers increased opportunity to meaningfully and structurally incorporate S&T teaching into their regular education programs; (2) *substantial* teacher professional development that helps primary teachers acquire the knowledge, skills, and positive attitudes to teach S&T by inquiry-based pedagogy and to alter their practices accordingly; and (3) *school-wide* implementations of inquiry-based pedagogy, in which *all* school members (including the school principals) are professionalized and involved with reforming daily school practice, which would likely foster leadership on school-wide policy, teacher collaboration and classroom experimentation, and shared positive culture for the adoption of inquiry-based pedagogy.

The report offered by the Exploratory Committee prompted the start of the current doctoral research project in the year 2013. Scientific descriptions of school-wide, inquiry-focused teacher professionalization have so far been scarce (Syer et al., 2012). This is not surprising, as large-scale and longitudinal school intervention studies are generally complex, expensive, and labor-intensive (Desimone, 2009). Partially because of this scarcity, little is known about what teacher and school factors might foster or hinder teachers' professional development in the adoption of S&T teaching (Shore, Aulls, & Delcourt, 2017; Thurlings, Evers, & Vermeulen, 2015). To help fill this void in the literature, this dissertation describes the results of a two-year school improvement program in which the complete school staffs of six Dutch primary schools were trained to integrate inquiry-focused (S&T) pedagogy into daily school practice. In the development of the program, the aforementioned main guidelines for the sustained implementation of S&T teaching were adopted. Below,

these main guidelines are described in more detail. Subsequently, we introduce the current thesis and provide an overview of the dissertation.

A broader focus on primary science and technology education

In many Western societies, policy documents promote education aimed at helping pupils become inquisitive, confident, and goal-driven young people who can solve the complex scientific and technological challenges of tomorrow and find meaning and pleasure in doing so (Pellegrino & Hilton, 2012). Such inquisitive, confident and goal-driven thinking is commonly associated with inquiry-focused competency (Walan et al., 2016). The implementation of inquiry-focused S&T education in primary education has therefore received increased global emphasis in the last decade (OECD, 2015; Osborne & Dillon, 2008).

However, operationalizations of inquiry-focused S&T education have undergone several advancements and have subsequently led policymakers, including the Exploratory Committee, to propose advanced notions about its implementation in primary education (Syer et al., 2012; Verkenningscommissie, 2013; Walan et al., 2016). Over time, inquiry-focused S&T teaching has been distinguished into the following three goals (e.g., Dewey, 1910; National Research Council, 2000, 2012; Osborne, 2014): (1) understanding how inquiry by scientists proceeds (i.e., learning *about* inquiry); (2) being able to perform inquiry (i.e., learning *to* inquire); and (3) constructing an understanding of (science) subject matter by inquiry (i.e., learning *by* inquiry).

Learning about inquiry

Attention for the value of inquiry-based pedagogy in primary education generally arose with the introduction of primary science curricula (Lumpe et al., 2000). Science and technology are regarded important subject areas for primary education, because economic, environmental, and societal challenges are becoming increasingly more scientific and technological in nature (Aikenhead, Orpwood, & Fensham, 2011; Potvin & Hasni, 2014). Primary education should thus help pupils become scientifically and technologically literate by teaching them key concepts in these domains, including an understanding of science-related skills, in order to participate in society (Ledoux et al., 2013; National Research Council, 2012; Osborne & Dillon, 2008).

To meet this goal, S&T-related lesson activities can involve metacognitive reflection and discussion, such as discussing with pupils the tentative nature of

scientific ideas, reflecting on the epistemic importance of inquiry for knowledge development and innovation in general, and considering their own potential roles as future scientists or engineers in society (Akerson, 2019; Deng, Chen, & Tsai, 2011; Khishfe & Abd-El-Khalick, 2002). As such, pupils come to learn about S&T as a *process* of inquiry and, thereby, learn how scientists and engineers may go about constructing explanations of natural phenomena or designing technology to solve problems.

Learning to inquire

While the first goal of inquiry in S&T education requires that inquiry be the focus of study, the second goal of inquiry requires that pupils learn to *participate* in such activities themselves – and not solely study, reflect on or discuss them. Scientific and technological innovations have become an integral part of everyday modern life (e.g., energy transition, climate change, globalization) and thus call upon more young scientists and engineers to contribute to these innovations at all levels of society (Levinsen & Nielsen, 2011). Therefore, helping pupils acquire basic inquiry competency in preparation of their professional lives is regarded increasingly important (e.g., OECD, 2015).

To achieve this, S&T-related lessons activities should familiarize pupils with the process of conducting inquiry, such as learning to formulate hypotheses, gather and interpret data, draw conclusions, and consider alternative answers or solutions to scientific questions and problems (Lederman, Antink & Bartos, 2014; National Research Council, 2012; Stender, Schwichow, & Zimmerman, 2018). Such activities mostly address the cognitive aspects of 'doing inquiry' and are often restricted to a predetermined science topic (e.g., magnets, buoyancy).

Learning by inquiry

Learning *about* inquiry and learning *to* inquire are distinct from *using* inquiry to learn content, although educators and researchers have been known to conflate these three goals (Hodson, 2014; Osborne, 2014). The third goal of inquiry, that of helping pupils *use* inquiry as a general strategy to study school subject matter, is tied to understandings of how individuals learn (Claxton, 2007). Research suggests that constructivist approaches to learning, such as inquiry-based learning, may help pupils develop a meaningful and integrated understanding of school subject matter, as they are actively involved in the construction of their own ideas, solutions and explanations (Bruner, 1961; Papert, 1980; Syer et al., 2012; Walan et al., 2016).

Based on this research, education policymakers have recently expanded their focus on the use of inquiry beyond the goals of primary science education. The core aspiration of recent international education policy is that being an inquiry-minded learner is useful in all facets of life and that, therefore, primary education should help pupils foster their inquiry thinking as an integral component of daily school practice (OECD, 2015). To that end, primary teachers should adopt inquiry-based pedagogy to prepare pupils for a lifetime of change: to teach them how to act when they are faced with questions, tasks or situations that are complex and for which they were not specifically prepared (Dede, 2010; Trilling & Fadel, 2009).

Private foundations and education organizations have since used a variety of labels for the sets of inquiry skills perceived as valuable in this respect, such as 'twenty-first century skills', 'key skills', or 'advanced skills' (e.g., Geisinger, 2016; Jerald, 2009; Pellegrino & Hilton, 2012; Voogt & Roblin, 2012). These skill sets are derived from a large research base in cognitive, developmental, social, and educational psychology and include such pupil qualities as curiosity, achievement-related motivation, and higher-order thinking skills. Notably, research views the skills associated with the labels "21st century skills" as reflecting fundamental dimensions of human competence that have been valued for many centuries, rather than qualities that are suddenly new, unique, and essential today (Logsdon, 2013).

From this broader view on S&T education at the primary level, inquiry is not just a means to familiarize pupils with scientific practice, nor a topic of mere reflection and discussion, but a means to help pupils mature into inquisitive, original and confident thinkers. As such, inquiry-based pedagogy is not limited to S&T content alone, but allows pupils' inquiry to take place across different subject domains, including traditional domains such as English, geography, and history (see Bennett, Lubben, & Hogarth, 2007; Heywood, Parker, & Jolley, 2012; OECD, 2015; Stuckey et al. 2013). It is believed that, through learning by inquiry, pupils develop a more meaningful and integrated understanding of school subject matter, including an understanding of the interrelatedness of S&T with traditional school subject matter, compared to forms of teacher-led instruction and rote learning (Osborne, 2014).

Changes in Dutch primary science and technology education
In line with the above, Dutch education policymakers have recently proposed broader views on the meaning of primary science and technology education as well (see Platform Onderwijs2032, 2016; Thijs, Fisser, & Van der Hoeven, 2014). In particular,

the Exploratory Committee proposed the adoption of the following broadened interpretation:

'Science and technology is a way of looking at the world. Science and technology education starts with wonderment: why is the world the way it is? From that attitude, questions spring or problems are identified. The search for answers to these questions and problems leads to solutions in the form of knowledge and/or products. These solutions are also the starting point for new questions.' (Verkenningscommissie, 2013, p. 6).

In the above terms, the committee characterizes 'S&T-minded' pupils predominantly by their inclination to be inquisitive. This perspective fits well with recent propositions in the international research literature on S&T education, as studies increasingly promote the importance of fostering pupils' curiosity to improve their academic achievement (e.g., Carr & Claxton, 2004; Engel, 2015; Grossnickle, 2016; Heywood et al., 2012; Jirout & Klahr, 2012).

Given this broadened perspective, S&T education should thus foremost encourage pupils' own inquisitive ideas and questions to emerge from their own study of all sorts of school topics and research projects. In turn, teachers should be 'responsive' to pupils' emergent ideas and questions by elaborating or extending these through group discussions, design experiments, or new investigations (Bennett et al., 2007). This way, inquiry stimulates pupils to curiously, creatively, and confidently study (novel) school subject matter (Claxton, 2007), to make productive connections between school topics (Heywood et al., 2012), and to develop a more meaningful understanding of the subject matter generally taught in school (Walan et al., 2016). This integral nature of S&T education would give pupils a better understanding of the interrelatedness of S&T with other school topics (such as literacy, geography, history and art) and at the same time offer primary teachers increased opportunity to teach S&T.

In this dissertation, we adopted the above focus of the Exploratory Committee in the development of the present school improvement program. Thus, we focused on inquiry pedagogy aimed at helping pupils *use* inquiry as a general strategy to study school subject matter. As such, the central aim of the program is explicitly linked to recent (inter)national policy on primary S&T education.

Substantial teacher professional development

Inquiry-based pedagogy expands teachers' use of teacher-led instruction with pupil-centered approaches to teaching that stimulate pupils' own investigation of school subject matter (Stuckey et al., 2013). It expands the role of the teacher to go beyond that of following (S&T) lessons chapter by chapter, to the role of enhancing lessons with opportunities for pupils to inquire, such as to design solutions to real-world problems, to conduct experiments, and to consider alternative answers or solutions to questions and problems (Claxon, 2007).

However, the complex task of changing the way inquiry is taught and assessed in primary school mostly falls upon the responsibility of the teachers to work out (Lumpe et al, 2000; Osborne & Dillon, 2008). This means that teachers should have answers to a wide range of questions about educational content, pedagogy, and how they should prepare themselves in these respects (see also Hargreaves & Fullan, 2012). Many of these questions still lack clear answers in the scientific literature today (Syer et al., 2012). It may not be surprising that most primary teachers thus struggle with adopting inquiry-based pedagogy (DiBiase & McDonald, 2015; Kim & Tan, 2011).

Helping primary teachers acquire the knowledge, skills, and positive attitudes to adopt inquiry-based pedagogy is thus a vital but complex enterprise, requiring substantial teacher professionalization (Capps, Crawford, & Constas, 2012). To our knowledge, few scientific descriptions of teacher preparation or enhancement programs exist that provide operationalized descriptions of the kinds of pupil qualities that inquiry-based teaching aims to develop and, subsequently, provide primary teachers an approach to pedagogy for developing and accessing these qualities in pupils (see also Claxton, 2007; Ledoux et al., 2013; Platform Onderwijsraad2032, 2016; Walan et al., 2016). In this dissertation, we thus used common themes from the scientific literature to define possibly relevant goals of inquiry-focused teacher professional development.

Positive attitudes towards inquiry teaching

Research indicates that teachers' inquiry teaching behavior is profoundly shaped by their attitudes towards inquiry teaching (Osborne et al., 2003; Lumpe, Czerniak, Haney, Beltyukova, 2012). Attitude can be understood as the evaluative beliefs a teacher may have about a particular behavior in a certain context in terms of favorable or unfavorable features (Ajzen, 2001), such as the benefit of the behavior or the pleasure of engaging in the behavior. These beliefs determine teacher's intention to

enact the behavior, either covertly or overtly, when sufficient opportunity arises to do so. Although attitudes are often regarded as stable personal beliefs that are difficult to change, attitudinal beliefs can be improved through intervention (Vogel & Wänke, 2016).

Most primary teachers find it difficult to teach school subject matter through inquiry, because they themselves lack sufficient familiarity with inquiry (Ricketts, 2014). As a consequence, primary teachers typically hold negative attitudes towards inquiry teaching and thus often shy away from such teaching (Jarvis & Pell, 2004). Encouraging inquiry teaching practice in primary education thus calls for teachers to become sufficiently familiar with enhancing opportunities for pupils to conduct inquiry and to improve their own attitudes toward inquiry teaching. Attitude training has therefore been regarded a vital component of inquiry-focused teacher professional development (Osborne & Dillon, 2008).

A recent experimental study by our research group showed that in-service primary teachers' attitudes towards S&T and towards inquiry teaching can be improved by six months of attitude-focused professional training (Van Aalderen-Smeets & Walma van der Molen, 2015). The training raised teachers' awareness about their own attitudes and challenged them to adopt more positive attitudinal beliefs where needed. Rather than providing prescribed science lesson examples, the focus of the training was thus mainly on realizing attitude change. In particular, teachers' improved self-efficacy beliefs (i.e., perceived competency to teach through inquiry) and context dependency beliefs (i.e., perceived dependency on available time and resources to teach through inquiry) showed to positively affect their frequency of inquiry teaching (Van Aalderen-Smeets & Walma van der Molen, 2015). For a detailed description of the theoretical underpinnings of these concepts, please see Van Aalderen-Smeets, Walma van der Molen, & Asma (2012). These results show that the improvement of teachers' self-efficacy and context dependency beliefs thus appears to be particularly relevant to the successful implementation of inquiry-focused teaching practice.

Lastly, we contend that inquiry-focused teacher professional development should aim at improving primary teachers' beliefs about creatively enhancing their usual teaching methods with inquiry teaching methodology (for theory about such beliefs, please see Thurlings et al., 2015). For example, teachers should feel inclined to revise teacher-led (science) lesson activities into more open-ended, student-centered lesson activities that allow pupils to conduct inquiry (Osborne, 2014). Most primary teachers rather prefer to follow lesson books chapter by chapter and, when

deciding to incorporate inquiry-based lesson activities, seek comfort in using prescribed inquiry lessons and materials instead (Jones & Eick, 2007). Inquiry-focused teacher professionalization should thus teach and motivate primary teachers to 'infuse' inquiry teaching methodology into their regular lesson programs.

Didactical knowledge and skills

Inquiry teaching practice not only requires teachers' *willingness* to encourage pupils' inquiry, but as much so requires their *ability* to do so. As previously described in this chapter, S&T education aims not merely at the transmission of content knowledge, but at teaching about the process of inquiry and to engage pupils in using inquiry to study school subject matter (Lederman et al., 2014; Osborne, 2014; Slavin, Lake, Hanley, & Thurston, 2014).

Such teaching methodology involves, among other, the implementation of minds-on and hands-on (science) lesson activities, which may include challenging pupils with conducting experiments, solving real-life problems, formulating open-ended questions about novel and complex subject matter, and adopting 'inquiry' as a fruitful strategy for learning rather than just a way that scientists conduct their work (Jarvis & Pell, 2004; Lederman & Abell, 2014; Miri, David & Uri, 2007; Osborne, 2014). This means that inquiry-based pedagogy is not confined to the teaching of complete, self-contained (S&T) lessons, but rather that such pedagogy should be integrated into the school curriculum by small lesson interventions (Claxton, 2007; Van Aalderen-Smeets, Walma van der Molen, Van Hest, & Poortman, 2017).

To help primary teachers develop such competency, professionalization activities should be aimed at helping teachers learn about the process of inquiry, the use of inquiry for school learning, and to familiarize them with what it means to learn by inquiry. Only if teachers acquaint themselves with what it means to be an 'inquiry-driven' learner will they be able to act as genuine 'inquiry-driven' role models to their pupils and engage them in inquiry (Akerson, Abd-El-Khalick, & Lederman, 2000).

In addition, teachers should learn to foster pupils' curiosity and wonderment throughout their practices, as this is widely promoted as one of the main objectives of S&T education in primary education (see Claxton, 2007; Jirout & Klahr, 2012; Verkenningscommissie, 2013). To that end, teachers should regularly convey to pupils the importance of pupils' own inquiries into school subject matter (Abd-El-Khalick, 2012) and confront them with novelty, unexpectedness, and uncertainty to elicit their curiosity and wonderment (Engel, 2015; Grossnickle, 2016). To further guide pupils' inquiry, teachers should use questioning techniques to foster pupils' higher-order

thinking (e.g., King, Goodson, & Rohani, 2011), encourage pupils to grow their inquiry ability by praising their efforts to persist their inquiry (e.g., Blackwell, Trzesniewski, & Dweck, 2007), and motivate them to achieve in school by promoting different achievement goals (e.g., Dweck, 2000). Simple reward systems could further emphasize to pupils that their inquiry behaviour is appreciated and part of the assessment of their overall learning in school (see also Claxton, 2007).

Extensive teacher professionalization

Based on the above review of the literature, we believe that a combination of attitude-focused and didactical training may thus provide primary teachers the necessary preparation for adopting inquiry-based pedagogy.

However, changing routines is difficult, especially when such change requires the mastery of new complex teaching skills (Desimone, 2009). Attitude change takes time as well, as teachers must reflect on and discuss their beliefs, gain positive classroom experiences, tackle their misconceptions where needed, and adopt positive beliefs (Osborne et al. 2003). Therefore, the Exploratory Committee states that inquiry-focused teacher professional development should be extensive enough to realize significant changes in school practice (Verkenningscommissie, 2013). Although schools likely prefer few, short and hands-on workshops on inquiry teaching, research indicates that such workshops will likely fall short.

Scientific guidelines for effective professional development indicate that professionalization should comprise at least a total of 50 contact hours and that significant effects can be expected over a minimum of two years (Borko, 2004; Desimone, 2009). During course meetings, participants should actively work on assignments and reflect on and discuss their learning experiences. They should also be prepared for take-home assignments to apply what they have learned from the meetings in their own classrooms. To accommodate this work, course meetings should be spread out over time to afford participants sufficient time for such implementation (and experimentation) in between meetings.

School-wide implementations of inquiry-based pedagogy

As described above, teachers play an integral role in determining the successful adoption of inquiry-based pedagogy. School practice reform can thus never be done *to* or *for* teachers, but only *by* and *with* teachers (Hargreaves & Fullan, 2012; Uiterwijk-Luijk, Krüger, & Volman, 2019). Teacher collaboration is therefore

considered an essential condition for successful school development (Kennedy, 2016). It provides teachers the opportunity to learn from and with one another by discussion, experimentation, and working towards shared goals (Hunzicker, 2011), which can be a powerful form of teacher learning (Desimone, 2009).

School principals are assumed to be important for setting favorable conditions for teachers to reform their practices in the aforementioned ways (Murphy & Seashore, 2018; Thoonen, Sleegers, Oort, Peetsma, & Geijsel, 2011). They should set and communicate clear policy on the direction and expectations of teachers' inquiry-focused teaching practice reform. By actively involving themselves with teachers' reform efforts and by learning from their challenges and advancements, they can formulate performance goals at the school level and decide on strategies to achieve these goals (for guidelines, please see Moolenaar, Sleegers, & Daly, 2012; Murphy & Seashore, 2018). They should further support and guide teachers' professional development, such as by encouraging them to experiment with new approaches to inquiry teaching and by facilitating teacher collaboration.

School principals with a strong inquiry-oriented vision about pupils' education, who prioritize school norms on classroom experimentation and teacher collaboration will likely be more successful with implementing S&T education than school principals who do not (Leithwood, Harris, & Hopkings, 2008; Thurlings et al., 2015). S&T-focused teacher professional development should therefore involve *all* school staff, including the school principals, to enable and empower leadership on creating favorable school conditions for sustained inquiry-based teaching practice.

The current thesis

In this dissertation, we adopted the above-described guidelines to develop a comprehensive teacher professionalization course, aimed at preparing in-service primary teachers to integrate inquiry-based pedagogy into their school curricula. In the development of the course, we thus explicitly adhered to (1) the broader focus on S&T education as recently promoted in (inter)national education policy documents, (2) the quality standards for teacher professional development aimed at helping primary teachers acquire the competency to teach (S&T) by inquiry-based pedagogy, and (3) a school-wide approach to the implementation of inquiry-focused teaching practice.

However, few scientific descriptions of inquiry-focused school improvement programs previously existed that provide operationalized descriptions of the school

principal, teacher, and pupil qualities that such programs should aim to develop. Therefore, as part of this dissertation, we needed to first define and operationalize many of these outcome variables ourselves.

Defining relevant outcome variables

At the school level, no validated measures previously existed of school principles' leadership behavior for fostering teachers' adoption of inquiry-based pedagogy, such as communicating clear school policy, facilitating teacher collaboration, and fostering positive cultural norms for practice reform. Similarly, at the teacher level, no validated instruments were available yet to assess teachers' inquiry teaching behavior, such as aimed at stimulating pupils' curiosity, inquiry ability beliefs, and their higher-order thinking. Therefore, based on the relevant literatures, we developed new measures that best exemplified the range of professional behaviors associated with each relevant aspect of school principals' leadership and teachers' inquiry-focused teaching practice.

In addition, we used the Dimensions of Attitude toward Science (DAS) questionnaire that was previously developed by our research group to assess teachers' professional attitudes towards science and inquiry teaching (see Van Aalderen-Smeets & Walma van der Molen, 2013). The questionnaire, and its underlying theoretical framework, provided operationalized descriptions of several relevant components of teachers' attitudes towards inquiry teaching, such as teachers' self-efficacy and context dependency beliefs. No validated measures previously existed to measure teachers' perceptions of creative, inquiry-enhanced lesson design. We thus newly developed this measure as well, based on the available literature.

Lastly, at the pupil level, we used existing guidelines and measures for operationalizing pupils' inquiry ability beliefs (i.e., perceived malleability of their inquiry ability), effort beliefs (i.e., perceived causality of their effort on achievement), achievement goal orientations (i.e., perceived goals for achieving in school), and higher-order thinking (i.e., the synthesis, evaluation or analysis of information in order to come to new solutions, ideas or questions).

However, few scientific guidelines and instruments previously existed to encourage and assess pupils' curiosity within school settings (Grossnickle, 2016). This scarcity is somewhat surprising, as the value of (epistemic) curiosity as a driver of complex and exploratory learning has been long promoted (Cook, Goodman, & Schulz, 2011; Loewenstein, 1994). Scientific descriptions of 'curiosity' in the literature have often been critiqued for confounding the concept of curiosity with the concepts

of interest, intelligence, or motivation, which makes it unclear what is measured or what mechanisms may underlie pupils' curiosity behavior (Silvia & Sanders, 2010). The few available studies on this topic have generally been limited to qualitative, explorative research (e.g., Engel, 2015) and suggest that pupils are provided little opportunity and encouragement to be curious in primary school, including during inquiry-focused lesson activities. In The Netherlands, however, pupils' curiosity has not yet been thoroughly investigated. In sum, these shortcomings were problematic in the development (and planned evaluation) of our school improvement program, because the encouragement of pupils' curiosity is promoted as one of the main objectives of primary science and technology education (e.g., Carr & Claxton, 2004; Jirout & Klahr, 2012; Verkenningscommissie, 2013). As part of this dissertation, we thus developed guidelines and measures for the stimulation and assessment of pupils' curiosity.

The intervention
Based on the above outcome variables, we developed a comprehensive inquiry-focused school intervention for six Dutch primary schools, which aimed at improving the knowledge, skills and attitudes of all school members for adopting inquiry-based pedagogy. To that end, the intervention included a nine-months teacher enhancement course. Because the positive development of teachers' attitudes was considered essential for successful practice reform, one of the main purposes of the course was to improve teachers' attitudes towards teaching inquiry by means of attitude-focused professional training. Therefore, the course included the complete six-months attitude-focused teacher professional development course by Van Aalderen-Smeets and Walma van der Molen (2015), which had proved to be successful in this regard.

In addition, we developed a consecutive three-month training course that aimed to help participants further familiarize with the broader concept of S&T teaching and learning (i.e., encouraging pupils to learn *by* inquiry), as previously described in this chapter. It also aimed at helping participants acquire the didactic knowledge, skills, and positive attitudes to foster pupils' curiosity, inquiry ability beliefs, effort beliefs, achievement goal motivations, and higher-order thinking as part of their daily practices. The participating school leaders did not receive specialized leadership training, but were required to participate in all course meetings and were encouraged to facilitate teachers' adoption of inquiry-based pedagogy during the program.

Research design

The above professionalization treatment was embedded in a two-year school improvement program, which allowed us to measure (changes in) pupils', teachers', and school principals' performance before, immediately after, and one year after the treatment. This longitudinal approach to evaluating the efficacy of the intervention was important, as the implementation of new school practices requires a minimum of two years (Desimone, 2009). The two-year time span of the program thus afforded us (and the participating schools) a realistic time span for significant practice reform.

A mixed-method approach to data collection was used to measure teachers' attitudes towards inquiry teaching, their inquiry teaching behavior, and school principals' leadership behavior (e.g., Likert-type questionnaires and semi-structured interviews). Likert-type questionnaires were used to measure the inquiry-related attitudes, beliefs, and motivations of the 4th, 5th and 6th grade pupils of the participating schools. Our main reason for focusing on this particular age group was that our survey instruments proved too difficult for younger pupils to comprehend.

Program effects where examined on the basis of a delayed treatment pretest-posttest control group design, which allowed all six primary schools to benefit from the training, while allowing us to compare the results of the treatment to school teams that did not (yet) receive the training. In addition, differences in the program effects between the individual schools were explored based on differences in school principals' leadership behavior. Therefore, we examined program effects at the treatment level (i.e., across the schools) and at the individual school level. This was important, because school improvement research indicates that intervention effects may likely differ among individual schools due to varying school leadership and organization (Berliner, 2002).

In sum, we believe that the above approach to researching the efficacy of our professionalization treatment would provide valuable insight into what professionalization features and what school and teacher factors may enhance the sustained implementation of inquiry-based pedagogy in primary schools. In addition, we aimed to contribute to the knowledge base of primary S&T education by offering (new) relevant and validated measures for assessing inquiry-focused school development. Pending good results, these measures could be used by other researchers to examine the effects of similar interventions.

Overview of the dissertation

This dissertation mainly describes the evaluation of the above school development program. As previously described, however, preliminary research indicated a scarcity of studies on the nature and dimensions of pupils' curiosity in the school context. Because pupils' curiosity was deemed a central objective of primary S&T education, and should therefore be included as an explicit measure for assessing school improvement in this field, separate theoretical and empirical research had to be conducted first. Therefore, the dissertation comprises two consecutive parts. The first two studies of the dissertation focus on the operationalization and measurement of pupils' curiosity. The last two studies focus on the evaluation of the school development program.

In the first study of this dissertation (chapter 2), we thus first set out to explore primary pupils' pre-existing concepts of, feelings towards and experiences with various types of 'curiosity' inside and outside the school context, across all grade level groups of two Dutch primary schools. This would bring insight into the aspects of pupils' curiosity that would require teachers' attention and, thereby, serve as an important stepping stone for the development of curiosity-focused pedagogy, teacher professionalization, and measurement instruments. These considerations were part of the second study of this dissertation (chapter 3). In the second study, we propose that pupils' curiosity can be understood in terms of their *images* of and *attitudes* towards curiosity. This attitudinal perspective on pupils' curiosity closely relates to the proposed definition of 'S&T-minded' pupils by the Exploratory Committee (Verkenningscommissie, 2013). We describe the development and validation of a novel instrument to assess relevant components of pupils' images of and attitudes towards curiosity, which we coined the Children's Images of and Attitudes Towards Curiosity (CIAC) questionnaire. As previously described in this chapter, we aimed to assess a broader range of pupil, teacher and school principal variables to evaluate the efficacy of our professionalization treatment. The development of these variables and their corresponding scales is not described in separate chapters of this dissertation, but included on a smaller scale as part of the third and fourth study (chapters 4 and 5).

The second half of the dissertation describes the effects of the intervention. In the third study (chapter 4), we describe the effects of the professionalization treatment on teachers' attitudes towards inquiry teaching, their creative lesson design beliefs, and inquiry teaching behavior during the two years of the school improvement

program. In addition, we describe changes in school principals' leadership behavior aimed at supporting and guiding teachers' adoption of inquiry-based pedagogy. Lastly, in the fourth study (chapter 5), we describe the extent to which teachers' changed inquiry teaching behavior during the program affected subsequent changes in their pupils' inquiry-related attitudes, beliefs, and motivations during this time. In addition, a Structural Equation Modeling approach is used to examine the relationships among pupils' attitude, belief, and motivation scores. Based on attitude and motivation theory, we investigate the extent to which pupils' attitudes towards curiosity and their implicit (inquiry) ability beliefs predict their efforts and motivations to be inquiry-driven learners in school.

In chapter 6, we conclude this dissertation by discussing the important findings of the overall study, its potential limitations, and recommendations for future research and practice.

In sum, this dissertation is based on four separate studies that are (or will be) published in scientific journals. Each study is self-contained, which means that each study includes its own theoretical introduction and discussion. All of the cited references in the main introduction and discussion of this dissertation, including the four studies, are jointly presented in a separate section ('References') at the end of this dissertation.

Chapter 2

2 Do children express curiosity at school? Exploring children's experiences of curiosity inside and outside the school context

Abstract

Education policies increasingly promote the development of children's epistemic curiosity in primary school as part of renewed (inter)national education standards. Yet, little is known about children's own conceptions and experiences of epistemic curiosity in school settings. In the present study, we interviewed 92 primary school children individually about their own beliefs, feelings, and accounts of curiosity inside and outside the school context. Results indicated that, at school, children barely experienced epistemic curiosity and generally perceived 'curiosity' as something that predominantly belongs to the social domain. Partly because of this narrow conception, the children did not attribute much learning-value to being curious in school and felt generally discouraged by their teachers to express their epistemic questions and ideas. However, many children reported to be actively curious about a diverse range of complex science topics outside of the school context. Based on these findings, we argue that curiosity-focused pedagogy should explicitly aim at cultivating a positive classroom climate in which children value the educational importance of posing epistemic questions and ideas, derive pleasure from being curious learners, and perceive that their teachers appreciate their curiosities. We conclude our paper with how such a positive classroom climate might be cultivated by teachers.

Introduction

In the last decade, the stimulation of children's epistemic curiosity in primary school has gained much attention (e.g., Engel, 2011; Jirout & Klahr, 2012; OECD, 2015; Pellegrino & Hilton, 2012; Spencer, Lucas, & Claxton, 2014). Epistemic curiosity is the desire to seek and acquire new intellectual information (Berlyne, 1954; Loewenstein, 1994; Piotrowski, Litman, & Valkenburg, 2014). International education policies increasingly promote the implementation of school curricula that aim to teach children about the epistemological importance of curious thinkers to society (Osborne & Dillon, 2008; National Research Council, 2012; Spencer et al., 2014). Such understanding is believed to entail not only factual knowledge about (scientific) discoveries made in the past, but also an understanding of the nature of knowledge-development itself and the social interaction that it requires (Fouad, Masters, & Akerson, 2015; Osborne & Dillon, 2008; Trevors, Muis, Pekrun, Sinatra, & Muijselaar, in press). To this end, education policy-makers increasingly call for investigative approaches to learning in primary school that engage children with discussions about knowledge-development or current socio-scientific issues. Such interactions may teach them about the tentative nature of scientific ideas and the epistemological importance of curious questions and ideas (Abd-El-Khalick, 2012; Kashdan, 2004; Lucas, Claxton, & Spencer, 2013).

In addition to fostering children's conceptions about the importance of curious question asking for the development of knowledge in general, researchers increasingly advocate the educational value of developing children's *own* epistemic curiosity (Baehr, 2013; Claxton, 2007; Claxton & Carr, 2004; Engel, 2011; Engel & Randall, 2009; Jirout & Klahr, 2012; Pellegrino & Hilton, 2012; Ritchhart, 2002; Tamdogon, 2006). Decades of developmental studies have shown that children's epistemic curiosity forms a key driver of their intellectual development (Chouinard, 2007; Cook, Goodman, & Schulz, 2011; David & Witryol, 1990; Kashdan & Roberts, 2004; Kashdan, Rose, & Fincham, 2004; Loewenstein, 1994; Piaget, 1952; Spielberger & Starr, 1994). Within educational settings, children's epistemic curiosity is associated with wonderment (Opdal, 2001; Pluck & Johnson, 2011), question-asking (Jirout, 2011; Jirout & Klahr, 2012), and explanation-seeking behavior (Arnone & Grabowsky, 1992; Berlyne, 1954; Litman, Hutchins, & Russon, 2005). Epistemic curiosity is believed to enhance children's persistence with learning (Metz, 2008; Simon, 2001; Von Stumm, Hell, & Chamorro-Premuzic, 2011) and to improve children's memorization of information (e.g., Gruber, Gelman, & Ranganath, 2014;

Hassan, Bashir, & Mussel, 2015; Jepma, Verdonschot, Van Steenbergen, Rombouts, & Nieuwenhuis, 2012; Kang et al., 2009). Thus, researchers suggest that primary education should not only aim at developing children's *understanding* of how knowledge is developed, but also at fostering their *willingness* to express and pursue their own epistemic questions and ideas about subject matter to improve their own learning (Lucas, Claxton, & Spencer, 2013; OECD, 2015; Pellegrino & Hilton, 2012; Spencer et al. 2014).

However, despite the seemingly widespread agreement on the importance of curiosity-eliciting educational content and pedagogy in primary schools, it seems that most primary school teachers devote little time to fostering children's curiosity (Engel, 2011, 2013; Engel & Randall, 2009). Research suggests that teachers often feel uncomfortable with stimulating children to express curious questions about topics that the teachers themselves often do not know the answers to (e.g., Van Aalderen-Smeets, Walma van der Molen, & Asma, 2012; Ramey-Gassert, Shroyer, & Staver, 1996; Ricketts, 2014; Schoon & Boone, 1998; Van Booven, 2015). Furthermore, in many countries, children are generally taught that there is just one correct answer to teachers' questions and that alternative explanation seeking – by being curiously minded and critically reflective – is disruptive to teacher-directed instruction (Claxton & Carr, 2004; Claxton, 2007; Rojas-Drummond et al., 2017). It seems likely that such everyday school practices will, over time, lead children to develop misconceptions about the educational value of being curious and may guide them away from their natural habit of questioning and exploring (Marx & Harris, 2006; McCombs, Daniels, & Perry, 2008). This reality is clearly not in line with the assumption that children's curiosity is vital to meaningful and complex learning and that, therefore, curiosity-eliciting learning activities should be made an integral part of children's education.

The question thus arises how we could bridge this gap between theory and practice. Unfortunately, scientific progress has been generally slow in this regard. While Maw and Maw (1964) were among the first to develop a measuring procedure for teachers to assess curiosity in children, it was only recently that researchers such as Jirout and Klahr (2012) and Engel (Engel, 2011, 2013; Engel & Randall, 2009) brought a renewed urgency to its scientific investigation (see also Luce & Hsi, 2014). For the last 60 years, curiosity research has been mostly negligent of the formal education context and focused primarily on the study and measurement of curiosity in adults. Or, when studies did concern children, mostly focused on their curiosity behavior in isolated or artificial laboratory settings (e.g., measuring the extent to

which a child investigates a particular toy). Most notably, there seems to be a lack of understanding of what children *themselves* think of curiosity. Many researchers and policy-makers have attempted to define curiosity *for* children, but curiously enough, in our review of the literature, we did not come across any studies that investigated children's *own* conceptions of what it means to be curious, either in or outside of the school setting.

In our view, these shortcomings hinder the effective development of curiosity-focused lesson content and pedagogy. Our lack of insight into children's general (mis)conceptions, feelings, and experiences of curiosity at school prevents us from classifying what aspects of their curiosity are generally underdeveloped and may thus hamper their potential curiosity engagement in the classroom. As has been well-established in the learning-sciences, in order for any educational change to occur, we should first understand children's pre-existing knowledge concepts and experiences about the topics or issues at hand, before we can effectively build-up their awareness, knowledge, skills, or attitudes (e.g., Bransford, Brown, & Cocking, 2000). In our view, this approach applies to the development of any pedagogy, and thus also to fostering children's curiosity.

Therefore, in the present study, we attempted to gain a better understanding of children's pre-existing concepts and experiences about 'being curious learners' at school and at home. Our goal ultimately is, of course, to design educational interventions and to set up teacher professionalization and we are aware of the importance of teacher-pupil and parent-child interactions in the development of children's epistemic curiosity. However, in order to effectively develop such curiosity-focused lesson content and pedagogy, for the present study, we deemed it necessary to focus specifically on children's *own* perceptions and experiences of curiosity.

In the following section, we first provide a brief overview of the main perspectives that have been postulated to describe the concept of 'curiosity' and we will touch on some of the methodological issues that have been raised to stimulate and measure children's curiosity behavior. Subsequently, we present the rationale of the present study and our research questions.

Defining, measuring, and promoting curiosity

Traditionally, curiosity is described in terms of behavioral characteristics. Berlyne (1954, 1960, 1978) was the first to classify four types of curiosity behavior: *perceptual* curiosity (i.e., aroused by novel visual, auditory, or tactile experiences and reduced by

exploration), *epistemic* curiosity (i.e., a desire for intellectual engagement or acquiring knowledge), *specific* curiosity (i.e., a desire for specific knowledge or information), and *diverse* curiosity (i.e., aroused by boredom or stimulation seeking). Berlyne's multidimensional view of curiosity received much attention in subsequent research (e.g., Byman, 2005; Kashdan, Gallagher, Silvia, Winterstein, Breen, Terhar, & Steger, 2009; Litman, 2008; Loewenstein, 1994) and encouraged others to characterize more 'specialized' curiosity behaviors, such as *scientific* curiosity (e.g., Jirout & Klahr, 2012), *information-seeking* curiosity (e.g., Litman & Spielberger, 2003), and *cognitive, physical* and *social* curiosity (e.g., Litman & Pezzo, 2007; Reio, Petrosko, Wiswell, & Thongsukmag, 2006).

These efforts have resulted in many curiosity behavior descriptions and related measurement instruments. However, many of these curiosity descriptions were later criticized for showing poor psychometric validity and reliability, containing too much conceptual overlap (Grossnickle, 2014), or requiring too demanding, complex or subjective measuring procedures (Mussel, 2010; Woo, Harms, & Kuncel, 2007). For instance, Silvia (2006) notes that behavioral observation measures of curiosity often show positive correlations with respondents' IQ levels or teachers' perceptions of students' intellectual status, rather than measuring curiosity per se. Furthermore, many scholars equate 'interest' with 'curiosity' and thus perceive curiosity as possessing cognitive, affective, and character variables (Ainley, 2006; Baehr, 2013; Kashdan & Silvia, 2009). Others have indicated that many self-report measures include item descriptions of states and traits of curiosity that are too abstract, such that respondents – especially children – find it difficult to understand and self-assess such descriptions (Chambers & Johnston, 2002; Jirout & Klahr, 2012).

In addition, no widely accepted conceptualization yet exists of what exactly causes children to be curious (Grossnickle, 2014). Berlyne (1954, 1960, 1978) suggested that curiosity could be best understood in terms of both state aspects (i.e., evoked by situational determinants) and trait aspects (i.e., relatively stable aspects that are explained by individual differences). Berlyne viewed curiosity as a psychological drive that is predominantly caused by environmental conflict (e.g., experiences of complexity, novelty, and surprise). Loewenstein (1994) suggested that curiosity is produced by unpleasant feelings of knowledge deprivation that motivate information-seeking behavior to diminish such feelings (see also Jirout & Klahr, 2012). Deci (1975), on the other hand, suggested that curiosity might be caused by the degree to which a person perceives himself or herself to be competent to bridge a particular knowledge gap.

In the past decade, scholars have argued the division of generally two types of curiosity, namely the distinction between interest-type curiosity (i.e., enjoying the acquisition of new information) and deprivation-type curiosity (i.e, feelings of relief when resolving unpleasant feelings of not-knowing) (Litman, Crowson, & Kolinski, 2010). Studies suggest that interest-type curiosity relates positively to mastery goal orientation motivations, while deprivation-type curiosity relates more to performance approach and avoidance orientations (Litman, 2008).

Based on the above descriptions, a variety of educational change projects have been suggested to stimulate the development and expression of children's curiosity. Some involved the use of child portfolios that require children to document their epistemic curiosities to enhance their awareness and curiosity about scientific topics over time (e.g., Jones & Shelton, 2011). Other studies suggest the organization of extra-curricular activities that introduce children to unfamiliar projects and topics that may broaden their interests, such as field trips, school exchanges, or company visits (e.g., Davidson, Passmore, & Anderson, 2010; DeWitt & Storksdieck, 2008; Post & Walma van der Molen, 2014). Engel (2011, 2013) proposes that curiosity-evoking lesson activities should foremost spring from teachers' own curiosity-driven role modeling to their pupils (see also Spektor-Levy, Baruch, & Mevarech, 2011). Similarly, Pluck and Johnson (2011) suggest that teachers should 'trigger' children's curiosity by confronting them with thought-provoking questions that make children aware of their own knowledge gaps.

In sum, the scientific literature on 'curiosity' presents a multitude of theories about the nature, determinants, and behavioral characteristics of curiosity. Although decades of research have clearly shed light on the complexity of 'curiosity', we agree with Jirout and Klahr (2012) that these efforts may also have steered us away from finding common ground. In our view, in order to find this common ground, we should first attend to the fact that thus far we have insufficient knowledge of children's *own* perceptions of curiosity, their curiosity experiences, and the potential learning-value that they adhere to being curious.

Present study

To fill the above void in research, in the present study, we made a first attempt to measure primary school children's conceptions of curiosity inside and outside the school context by means of a structured interview procedure. Because of the exploratory nature of the study, we formulated research questions, rather than hypotheses. Research questions of interest were: In what way do children understand

what it means to be 'curious'? How do they describe their feelings of being curious? What do they believe is the relevance of being curiously minded? To what extent do they recognize their own curiosity experiences when prompted by different types of curiosity behavior? And what differences might exist between children's responses when we consider context (inside or outside the school context) or children's age?

To investigate children's own curiosity experiences, we prompted them with examples about: (a) *sensory* curiosity (e.g., wanting to know the origins of novel or sudden sounds, determining the particular taste of novel food, etc.); (b) *cognitive* curiosity (e.g., wanting to know how computers work, where babies come from, whether there is intelligent life on other planets, etc.); (c) *epistemological* curiosity (e.g., wanting to know how television was invented, how medicines were developed, how electricity was discovered, etc.); and (d) *wonderment*, which we defined as consciously noticing or 'being struck by' everyday particularities that seem valuable or meaningful (e.g., noticing the way tree leaves show many different vibrant colors from season to season, being struck by the way that birds fly, feeling perplexed about the complexity of new technologies, etc.).

We selected these four types of curiosity behavior in close conjunction with commonly used descriptions of the 'curious learner' in current international education policy documents (e.g., Lucas et al., 2013; OECD, 2015). The educational value of children's *sensory* and *cognitive* curiosity is widely recognized and believed to be especially prominent in young children's exploratory behavior (Berlyne, 1960; Kashdan & Steger, 2007). *Epistemological* curiosity, on the other hand, only recently gained more attention by education policy-makers as a result of the renewal of science curricula that specifically aim to engage children in the process of knowledge development (Olson & Loucks-Horsley, 2000; Tai, Qi Liu, Maltese, & Fan, 2006). Lastly, the assumed educational value of *wonderment* – as a possible precursor or after effect of exploratory curiosity – is often referred to by educators in non-academic work, but has hardly ever been researched before (Opdal, 2001).

Method

Participants

Two Dutch primary schools from medium-sized towns participated in the study. From both schools, 4 boys and 4 girls were randomly selected from Grade 1 through Grade 6 to be individually interviewed by the principal researcher. Four child interviews were later excluded from our dataset because they were found to be largely

incomplete, due to disruptions of the interview by parents or teachers. Thus, the total number of child interviews that we examined was 92 (46 boys and 46 girls; please see Table 2.1 for the number of boys and girls per grade level that participated in the study). Teachers and children were not made aware of the goals of the interview but were debriefed after the study was completed.

Table 2.1

The number of boys and girls per grade level that participated in the study

Grade level	Boys	Girls	Total
Grade 1	8	8	16
Grade 2	8	8	16
Grade 3	8	8	16
Grade 4	8	8	16
Grade 5	7	7	14
Grade 6	7	7	14
Total	46	46	92

Interview measures

The principal researcher held structured interviews with each individual child. We used a standardized format of open-ended questions that were consistently repeated to all children. At the beginning of the interview, children were asked to provide their age and gender. The next section of the interview consisted of four consecutive subdivisions with open-ended questions that measured: (1) Children's personal description of curiosity (i.e., 'Can you explain what curiosity means? Please give us your own description'), (2) Examples of general self-reported curiosities and related feelings (i.e., 'What are you usually curious about? And how do you feel in that particular case?'), (3) Examples of self-reported curiosities and related feelings, specific to the school context (i.e., 'What are you usually curious about at school, during class? And how do you feel in that particular case?'), (4) Children's perceived relevance of being curious (i.e., 'Do you think that it is important for people to be curious? Please explain your answer'). In this section, children were allowed to provide as many answers as came to mind.

In the second part of the interview, we asked children to share examples of their 'curiosities' separately for each of our proposed curiosity behavior dimensions (i.e., sensory curiosity, cognitive curiosity, epistemological curiosity, and wonderment).

For each dimension, we provided two to three curiosity behavior questions that were related to potential, everyday curiosity experiences. For each curiosity behavior question, we first prompted children with a context-free question that asked them to share up to two personal curiosities that first came to mind and we next prompted them with a context-specific question that explicitly asked them to share up to two curiosities that were specific to their time at school (i.e., a child could provide a maximum number of four curiosity accounts per curiosity behavior question). This approach allowed the children to share school-specific curiosities also in response to our initial context-free question prompts (for each curiosity behavior dimension), if such accounts would first come to mind. In this way, we intended to gain additional insight into the degree to which children implicitly associated each type of curiosity behavior with their time in school, prior to asking them explicitly to share any school-specific accounts. In this second part of the interview, we limited the number of possible accounts that children could provide here to two for each context (i.e., context-free and school-specific prompts), in order to keep the administration time short and to avoid taxing children's attention span too much.

In this structured manner, we derived children's accounts of: (1) Sensory curiosities (i.e., 'Do you like to explore the environment? If so, please provide an example', 'If you experience something that is unfamiliar to you, would you like to know what it is? If so, please provide an example', and 'Do you enjoy experiencing novel things? If so, please provide an example'), (2) Cognitive curiosities (i.e., 'Do you consider yourself a questioner? If so, please provide an example', 'Do you like to find out about how things work? If so, please provide an example', and 'Do you have follow-up questions when you find out about something new? If so, please provide an example'), (3) Epistemological curiosities (i.e., 'When you find out about something new, do you like to know how someone discovered that knowledge? If so, please provide an example' and 'Are you interested to know how people made certain inventions? If so, please provide an example'), and (4) Wonderment (i.e., 'Do you sometimes suddenly notice something particular about ordinary things? If so, please provide an example', 'Do you sometimes suddenly notice something particularly interesting in your surroundings? If so, please provide an example' and 'Do you enjoy noticing things that are special? If so, please provide an example').

Procedure

Children were individually interviewed during school time in a separate room outside of the classroom. The researcher used a pre-structured paper format to administer

children's responses. When children experienced difficulty answering a particular question, the researcher provided encouragement only once by providing a possible answer example before continuing the interview (when children copied these responses as their own answers, they were coded as 'copies of encouragements'). In some cases, when it was unclear whether a child's personal account of curiosity was related or unrelated to the school context, the researcher asked the child to clarify. We kept the administration time to a minimum by alternating between two versions of the interview. Both versions contained all of the above-described questions about age and gender, children's description of curiosity, their self-reported general curiosities, their self-reported school-related curiosities, and their perception of the relevance of being curious. The two versions differed in the types of curiosity behavior dimensions that were covered in the second part of the interview. One version only covered the dimensions of sensory curiosity and cognitive curiosity, while the other version covered epistemological curiosity and wonderment. These alternate versions of the interview were administered randomly among boys and girls for each grade level, which resulted in 47 children who were asked about sensory and cognitive curiosity and 45 children who were questioned about epistemological curiosity and wonderment. This way, the average duration of the interview was kept to about 10 minutes. Children were presented with a small gift after they had completed the interview.

Coding and scoring

The principal researcher and a second assessor coded all children's responses. For each interview question, responses were analyzed and categorized into a large set of detailed categories that were based on distinct 'domains' of curiosity examples that emerged from the data. These specific categories were later aggregated to create more general, higher-order categories by combining closely related categories. Both assessors then independently coded children's responses by marking categories as either applicable or not. If a response could not be categorized, both the principal researcher and the second assessor reviewed the response and decided if a new code should be added or if it should be put under the Other category. Approximately five new codes were added in this manner during the coding process. Inter-rater agreement was calculated using Cohen's kappa for ten per cent of the total number of child interviews. Overall inter-coder reliability was found to be good, as indicated by reliability levels that ranged between .89 and .93 across all sections of the interview. Our final set of categories, along with the response frequencies and types of curiosity

examples given for each category, are reported for each interview section in the Results section below.

Results

Children's responses are organized below in the same structured format that we used to administer the interviews. To investigate potential differences between children's responses from different grades, we started off with grouping children into lower (Grade 1 and 2), middle (Grade 3 and 4), and upper grade levels (Grade 5 and 6). However, we observed that children's responses were very much alike across these different grade levels, with respect to our research objectives: (1) children's curiosity definitions and their perceived relevance of curiosity were largely similar; (2) most children shared a similar number of curiosity accounts in response to all of our context-free curiosity question prompts for each curiosity behavior dimension; and (3) while children mostly shared personal curiosities from their time *outside of school*, they barely shared any personal curiosities related to their time *in school* in response to both our context-free and school-specific question prompts for all curiosity behavior dimensions. Because we observed no differences across age levels in children's responses in these respects, we disregarded the group variable from subsequent result descriptions.

Children's definitions of curiosity

We started the interview with asking children to define what it means to be curious ('Can you explain what curiosity means? Please give us your own description'). In total, 91 children were able to provide a definition (one child in the first grade was unable to do so). Most children framed their curiosity definitions as a *desire to know*, but differed in the objects of curiosity that they described. Based on these differences, we derived the following four categories that best fitted children's curiosity definitions: (a) *Novelty*, a desire to get to know something new (46%); (b) *Gossip*, a desire to know a rumor about someone else (24%); (c) Secret, a desire to know someone's secret (7%); (d) *Planning*, a desire to know what someone has planned for me/us to do next (3%); the remaining number of curiosity descriptions were categorized as other (20%) and concerned ambiguous responses (e.g., 'Feeling excited', 'To hear something', etc.). This result reveals that, while almost half of the children defined 'curiosity' as a general desire for new information, about a third of

the children associated curiosity specifically within the social context, in particular, to acquire private or secret information about others.

Children's curiosity accounts

The next section of the interview invited children to share any personal curiosities from their everyday lives that first came to mind ('What are you usually curious about?'). Children were allowed to share as many curiosities as they could. Once children had no more curiosities to share, we asked them to describe their feelings for each reported case ('And how did you feel in that particular case?').

In total, children initially provided 120 personal accounts of curiosity, indicating that on average children provided at least one account each (M = 1.30, SD = .53). Most curiosities that were shared in response to our first context-free question were related to the context of children's everyday lives outside of school (83%). The remaining 17% of accounts were school-specific. Only 13 children (14%) spontaneously shared curiosities from both contexts. When we subsequently asked the children to share any curiosities specific to formal lesson activities at school ('What are you usually curious about at school, during class?'), 68 (74%) children shared an additional 76 curiosities. This result reveals that, while children reported relatively few school-specific curiosities in response to our initial context-free question, children did report a greater number of school-specific curiosities when they were prompted to do so explicitly.

Overall, in response to our initial context-free question, most children reported being curious about the *Gifts* that they would receive for their birthday or for Christmas (33%). This particular curiosity was closely followed by curiosities about the *Planning* (28%) of upcoming family trips or lesson activities; *Eavesdropping* (26%) on private conversations between friends, classmates, or parents; and other trivial examples (13%). When we asked children to share the contents of their curiosities specifically related to formal lesson activities, most reported curiosities remained related to getting to know the teacher's lesson *Planning* (41%); followed by seeking for the teacher's *Help* (20%) with regard to operational questions, such as how to complete a math assignment or spell a word; *Eavesdropping* (17%) on social gossip between classmates during class; anticipating one's test *Score* (11%); and other trivial examples (12%).

These results show that, although children did report a fair number of personal curiosities related to the school context, most of these accounts seem to have little meaningful relation to investigative learning. They rather concern more trivial

curiosities about lesson planning, operational tasks, or obtaining private or secret information about classmates. We were particularly struck by the fact that a few children responded quite surprised or even disturbed when we asked them to share their school-specific curiosities: 'No one is curious about what we learn in class. We just need to do whatever the teachers tell us to do', 'No, of course not. It does not matter whether I am curious, because we just need to learn whatever we are assigned to do', 'Are you joking? There is nothing to be curious about, when doing boring math or reading'.

We continued by asking them to describe, for each reported case separately, how they recalled 'their feeling of being curious'. Irrespective of context, all children consistently described their accompanying curiosity feelings as mixed feelings of excitement and frustration. Children elaborated on their responses by explaining that, while it may feel exciting to get to know something new and interesting, it simultaneously feels frustrating that someone else is already knowledgeable of the desired information that 'you' have been excluded from. Other children specifically described curiosity as feeling 'nosy' or 'intrusive', and reported that others (e.g., classmates, teachers, friends or family members) explicitly discouraged them from poking their nose in certain matters. These responses reveal that, although children's accompanying feelings of being curious strongly relate to deficit-type definitions of curiosity in the literature (i.e., as a psychological drive that is caused by unpleasant feelings of not- knowing), their conception of curiosity predominantly relates to 'being nosy about personal matters', rather than to other types of knowledge exploration.

Perceived relevance of curiosity
We continued the interview by asking children to what extent they believed curiosity is valuable to people's lives ('Do you think that it is important for people to be curious? Please explain your answer'). Eight lower graders and one middle grader were unable to provide a sufficient answer, suggesting that this particular question might have been too difficult for some of the younger children to understand. For those children who were able to respond, we derived the following four categories that best fitted their answers: 42% of the children responded with 'Yes it is important, because curiosity makes us discover new things about others'; 22% responded with 'Yes it is important, because being curious about others makes life exciting'; 17% responded with 'No curiosity is not relevant, because a lot of things should remain private'; 13% responded with 'No it is not important, because we will eventually come to know

things about the world through the teachings of others anyway'; the remaining 6% of the children provided miscellaneous other answers.

These results again suggest that the children in our sample predominantly perceived curiosity as something that occurs within the social context (i.e., being nosy, wanting to know information about others). While this led some children to reason that people's curiosity should in fact be encouraged because it makes you learn more about other people (64% in total), other children believed that the development of curiosity should be discouraged (17%), because being curious about other people's private lives is often considered intrusive and disrespectful. It should be noted that the remaining 13% of the children did consider the potential value of curiosity in a broader, educational context. However, these children mostly disregarded the educational value of curiosity, because they believed that they would come to learn about the world through the teachings of others anyway, irrespective of whether they would be curiously engaged or not. Only one child (upper grade level) answered that curiosity is important because it may lead to new inventions.

All in all, the results on our question about the perceived relevance of curiosity clearly show that most children (from every age group) held naïve notions or misconceptions about what it means to be curious and did not seem to connect the concept of curiosity to education.

Children's personal curiosities for each separate curiosity behavior dimension

In the last section of the interview, we invited children to share their personal curiosities in response to each of our proposed curiosity behavior dimensions (i.e., sensory curiosity, cognitive curiosity, epistemological curiosity, and wonderment). Because we observed very similar patterns in children's responses across all four categories, below, we only outline our overall findings. Please see Table 2.2 for an overview of the number and kind of responses to our context-free and school-specific questions, for each separate curiosity category and related curiosity behavior question prompts.

Table 2.2

Children's curiosity accounts in response to each separate curiosity category and related curiosity behavior question for both context-free and school-specific prompts

Curiosity category	Context-free						School-specific		
	Outside of school			Inside of school			Inside of school		
	N	n	Response categories	N	n	Response categories	N	n	Response categories
Sensory curiosity									
Do you like to explore the environment?	44	47	Abroad, Local, Accidents, Sound, other	1	1	Sounds	30	30	Local, Accidents, Sounds
If you experience something that is unfamiliar to you, would you like to know what it is?	41	49	Sounds, Animals, Tastes, Eavesdropping, Words, other	2	2	Sounds, Words	23	24	Eavesdropping, Sounds, Words, Animals, other
Do you enjoy experiencing new things?	46	52	Trips, Holidays, other	4	7	Books, other	28	29	Books, other
Cognitive curiosity									
Do you consider yourself a questioner?	33	35	Eavesdropping, Help, Planning, other	3	6	Help	32	33	Help, Planning, Eavesdropping
Do you like to find out about how things work?	45	57	Computers, Vehicles, Biology, Astronomy, Society, other	2	2	Help	26	27	Help, Interactive Whiteboard, other
Do you have follow-up questions when you find out about something new?	24	25	Computers, Vehicles, Help, Eavesdropping, other	7	7	Help, other	16	16	Help, other

Table 2.2 continued

Epistemological curiosity

When you find out about something new, do you like to know how someone discovered that knowledge?	31	34	Eavesdropping, other	1	1	Eavesdropping	11	11	Teacher, other
Are you interested to know how people made certain inventions?	42	50	Computers, Machines, Physics, Astronomy, other	0	0	n.a.	27	29	Math, Computers, Machines, other

Wonderment

Do you sometimes suddenly notice something particular about ordinary things?	25	27	Objects, Structures, other	5	8	Objects, Structures, other	19	19	Objects, Structures, other
Do you sometimes suddenly notice something particularly interesting in your surroundings?	36	36	Abroad, Local, Instructions, other	2	3	Instructions, Accidents, other	24	25	Instructions, Animals, Objects, other.
Do you enjoy noticing things that are special?	26	28	Funny, Danger, other	2	2	Danger	12	12	Funny, Schoolyard

Note:
N is the total number of children who provided an answer with respect to the particular response group; *n* is the total number of curiosity accounts shared by the children with respect to the particular response group; the response categories are ordered from most to least often reported by the children.

More than half of the children were able to report at least one curiosity example for each of our curiosity behavior questions. In response to our context-free prompts, for all curiosity behavior dimensions, most curiosities that were shared by the children related to their everyday experiences outside of the school setting (91%). Only 9% of children's spontaneous responses were school-specific. In line with our earlier observations, children reported a greater (but still marginal) number of school-specific curiosities when we continued with specifically asking them about their curiosities in the school context. In line with several responses to our earlier interview questions, some children were surprised or even distressed when we asked for school-specific curiosities, as according to these children, everyday lesson activities in school 'obviously' did not stimulate their curiosity, let alone, allow for room to further pursue them. This result again shows that, irrespective of which one of the four curiosities we asked children to relate to, children associated their curiosity experiences mostly with their lives outside the school setting.

Our results showed that most reported *school-specific* curiosities, either in response to our context-free or school-specific prompts, were unrelated to investigative, explorative learning. For example, in case of children's sensory curiosities, children mostly reported being curious about the introduction of new lesson books because they wanted to browse for new and attractive images, or that they were curious about sudden sounds coming from neighboring classrooms. In the case of cognitive curiosities, children reported being curious about the correct answers to assignments and teachers' help in this respect, figuring out the teacher's lesson planning, eavesdropping on private conversations between classmates, or how the touch screen of the interactive whiteboard works. Children indicated that these curiosities did not originate from any formal lesson activities, and that their teacher had not acknowledged their curiosities nor provided room to further pursue them as part of class or a later project. In case of children's epistemological curiosities, children reported being curious about the invention of math, computers and, strikingly, how the teacher could have become so knowledgeable about all sorts of things in the world. Finally, children's responses to our wonderment behavior prompts revealed only trivial examples that were related to the school setting, such as children noticing the use of unfamiliar instructions or words by the teacher and children spotting new desktop accessories or printers in their classroom.

Contrary to the above-described results, children did report a fair number of investigative, explorative curiosities that were related to their lives outside the school setting. Especially in the case of cognitive and epistemological curiosities, children

shared many instances where they used the Internet (e.g., Wikipedia, YouTube, etc.) at home to learn about complex topics, such as computers (e.g., 'Where does the Internet come from?'), physics (e.g., 'What is electricity?'), astronomy (e.g., 'Why are planets always round?'), biology (e.g., 'Where do babies come from?'), machines (e.g., 'How does a microwave work?), and society (e.g., 'Who pays for the workers that build new roads?'). These results reveal that, while children hardly reported any curiosities relevant to learning that were related to the school context, outside of school, they apparently experienced a far greater number of explorative curiosities about complex topics that seem highly relevant to formal learning.

Discussion

We set out to explore children's own notions of what it means to be curious. We used a structured interview procedure to ask 92 primary school children about their own beliefs, feelings, and expressions of curiosity inside and outside the school context. As far as we know, these aspects of children's curiosity have not been investigated before. In our view, this lack of understanding has thus far prevented us from examining what aspects of children's curiosity may require particular attention in curiosity-eliciting content and pedagogy. Below, we outline and discuss the main findings of our study. In the second part of this Discussion, we argue that curiosity-focused pedagogy should not only aim at developing children's *skills* to pose epistemic questions and ideas in school, but should also aim at cultivating a positive classroom climate in which children value the educational importance of posing questions and ideas, derive pleasure from being curious learners, and perceive that their epistemic questions and ideas are appreciated by their teachers. We conclude our paper with how such a positive classroom climate may be cultivated by teachers.

Main findings of the present study

   Children's reported curiosity at school. Overall, did we find that children experience curiosity at school? Yes, we did to some extent, if we go along with children's reported conception of curiosity as something that predominantly concerns being curious about other people's private affairs or about lesson planning. While such a concept of curiosity does form an important aspect of children's general development (e.g., Dewey, 1910), it does not necessarily relate to the sense of epistemic wonderment and interest that we would like to stimulate in a formal school setting (e.g., Arango-Muñoz, 2014; Claxton & Carr, 2004; Engel, 2009, 2011; Jirout &

Klahr, 2012; Kashdan & Steger, 2007). Irrespective of children's age or the types of curiosity behavior that we asked them to respond to, the children in our sample reported few spontaneous examples of epistemic questions that related to the formal school setting. When children did recount examples of curiosity at school, these were almost exclusively directed at social issues, such as eavesdropping, or generic school-related issues, such as lesson planning or wanting to know an answer to a math problem.

Children's conception of curiosity – as something that predominantly belongs to the social domain – also characterized their reported feelings of curiosity. Children consistently described their curiosity feelings to be a mix of excitement and unpleasantness. To some extent, these reported feelings of interpersonal curiosity resemble the deficit-type descriptions of curiosity as suggested by others (e.g., Litman et al., 2010; Litman & Pezzo, 2007). The children in our sample on the one hand felt interested in 'getting in on' the private affairs of others, while on the other hand felt socially 'excluded from' and 'insecure about' not having obtained the information earlier. Some children further explained that peers and family often discouraged them to 'stick their nose into private affairs, leading them to perceive 'curiosity' as a particular behavior that others often disapprove of. Studies in the domains of social and developmental psychology suggest that this particular focus on social curiosity during childhood is not related to possible cultural differences in pedagogy, but seems to be a cross-cultural phenomenon (e.g., Dewey, 1910; Grossnickle, 2014; Litman & Pezzo, 2007) that may facilitate the forming of friendships (e.g., Rosnow, 2001) or prevention of social confrontations (e.g., Galen & Underwood, 1997).

Nevertheless, for educational purposes these findings are discomforting because they show that the children in our sample held rather narrow and naive conceptions about the meaning and use of being curious. Although not empirically tested before, our results do confirm observations by others. For example, Engel (2009, 2011) attempted to investigate what kinds of curiosity children express at school but was surprised by the lack of curiosity expressed. In line with Engel, our findings suggest that at school most children (a) barely engage in epistemic curiosity, (b) generally perceive their curiosity behavior as something that their teachers (or their peers or family members) disapprove of, and (c) attribute little value to curiosity as a learning strategy. Overall, children reported only few meaningful school-specific curiosities and, in some cases, even responded surprised or distressed about the possibility of such curiosity to be allowed in educational settings in the first place.

Although discomforting, our results are highly relevant. They suggest first that many children may hold misconceptions about the nature and value of curiosity as a driver for learning. Second, our results suggest that teachers seem largely inattentive to children's curious questions, fail to make children aware of the educational and scientific value of curiosity, and do not seem to act themselves as curiosity-eliciting role-models to their students. Lastly, our results suggest that children predominantly perceive curiosity as a desire for new information about other people, which is referred to as interpersonal curiosity (Litman & Pezzo, 2007).

Children's reported curiosity outside of school. Does this mean that children experience no meaningful curiosities at all? No, not entirely. The remaining results of our study indicated that about half of the children did report being curious about a diverse range of complex (mainly scientific and technological) topics in settings *outside* of school, especially in response to our cognitive and epistemological question prompts. For example, children reported being curious to learn about the governmental organization of society, the discovery of electricity, dinosaurs and the planets, the invention of computers, car engines or refrigerators, etc.

Given this result, it seems to us that we do not necessarily need to teach children how to *be* curious. This finding is promising and confirms once more the well-established notion that many children are naturally curious about many complex (scientific) phenomena, even without the use of formal education (Berlyne, 1960; Kashdan & Steger, 2007; Piaget, 1952). At the same time, however, our findings suggest that the lack of curiosity expressed by children in the classroom is not so much because children's inquiry skills fall short, but because our schools' social norms, beliefs, reward systems, and pedagogy insufficiently stimulate children's curious expressions in the classroom.

As part of the debriefing, we presented our main findings to the teachers of the children who participated in the current study. Most teachers were distressed to learn about their pupils' reported lack of cognitive and epistemological curiosity experiences in their schools. Moreover, many teachers expressed their concern about some of our children's reports in which they recounted that their teachers provided them little opportunity to reveal their cognitive and epistemological curiosities, and that some teachers even discouraged them to do so. However, many teachers also emphasized the educational importance of pupils' epistemic curiosity and that they believed to provide sufficient inquiry opportunities to curiously explore and study subject matter.

In many countries, however, inquiry-based teaching often reflects a 'learning *about* inquiry by *doing* inquiry' approach (Lederman, Antink & Bartos, 2014). In these terms inquiry is interpreted by many teachers as involving a "mechanical" process of cyclical inquiry activities, such as formulating hypotheses, gathering and interpreting data, and drawing conclusions (National Research Council, 2012). Such an interpretation of inquiry addresses the *cognitive* aspects of inquiry, but often overlooks the *metacognitive* aspects of inquiry that are deemed equally important to children's inquiry performance (Deng, Chen, & Tsai, 2011; Khishfe & Abd-El-Khalick, 2002), such as addressing the epistemic use of curiosity as an important metacognitive strategy to discover useful subject matter or to come to new ideas (e.g., Cook, Goodman, & Schulz, 2011; Von Stumm, Hell, & Chamorro-Premuzic, 2011). In fact, observations by Claxton (2007) and Engel (2011) suggest that many primary school teachers seem to perceive children's curiosity to be a natural 'byproduct' of inquiry and thus believe that the stimulation of children's curiosity does not require their explicit attention. Our current findings suggest that this perception may be false. Furthermore, research indicates that this omission may even help to explain why inquiry-focused teaching often fails to engage children in authentic inquiry and, consequently, fails to lead to improved knowledge attainment (Hodson, 2014). When left to their own discoveries, children may likely not interpret teachers' prompts for inquiry as opportunities for curious thinking and thus not achieve the kind of curiosity-driven inquiry that teachers seek.

Promoting children's curiosity in the classroom

To remedy this problem, we believe that teachers should adopt a more explicit approach to teaching children *about* curiosity, alongside inquiry-focused lesson activities. Prior research shows that inquiry can foster changes in children's views about inquiry when used together with explicit reflection and discussion activities (Abd-El-Khalick, 2012); Deng et al., 2011; Southerland, Johnston, & Sowell, 2006).

Through reflective group discussions, for example, teachers can help make children become aware of their predominant social concepts of curiosity and, subsequently, inspire them to adopt more epistemic images of curiosity (e.g., Abd-El-Khalick, 2012). Such discussions can either be held as stand-alone interventions or held in relation to particular curiosity-eliciting subject matter or project assignments (Deng et al., 2011). Both types of discussion can be understood as metacognitive activities that stimulate children to consider their existing concepts and experiences of curiosity, to compare these with more epistemic notions of curiosity, to discuss the

possible uses of curiosity for their own learning, and to learn to recognize and articulate their epistemic wonderments and questions. In parallel, teachers may guide such group discussions by illustrating the ways that epistemic curiosity drives pupils (and professionals) to explore new subject matter, to creatively come to new ideas or alternative explanations to existing problems, and to improve their overall academic achievement. Such reflective group discussions may help to broaden children's naïve and narrow notions of curiosity and may offer teachers opportunities to achieve sufficient 'cognitive dissonance' among children in order to re-orient their perspectives of curiosity (Bricker & Bell, 2008; Piaget, 1969) and to build-up their epistemic question-asking skills. However, this does not mean that children's social curiosity has no place in formal education. For example, teachers could deliberately employ children's social curiosity to foster forms of collaborative discovery learning (e.g., what do you think your classmates are thinking about a certain problem or assignment) or to use their curiosity for other people's thinking as a means to understand knowledge development in general (e.g., what do you think a scientist might have been thinking when he/she invented a new solution to a certain problem).

In addition, we strongly agree with Engel (2011) that teachers need to adopt a positive attitude toward epistemic curiosity *themselves* and need to experience first-hand why it is important and what it feels like to pose creative, explorative questions about any subject matter at hand (see also Tsai, 2006; Spektor-Levy et al., 2011). Studies show that children's views about inquiry and their learning goals are strongly influenced by teachers' views (Osborne, Simons, & Collins, 2003). Only when teachers genuinely act as curiosity-driven role models to their pupils will they be able to instigate a sense of what it means to be a curious learner in their pupils (e.g, Akerson, Abd-El- Khalick, & Lederman, 2000). They should pay explicit attention to cultivating and maintaining positive social classroom norms throughout their teaching in order to make children feel that their epistemic wonderments, questions, and ideas are appreciated (Amabile & Pillemer, 2010). Simple reward systems could further emphasize to children that their epistemic curiosities are very much encouraged and part of the assessment of children's overall learning in school (see also Teo, 2013).

Lastly, research suggests that children's scientific interests and ways of sense-making at home can be used as academic resources in teaching and learning at school (e.g., Warren, Ballenger, Ogonowski, Rosebery, & Hudicourt-Barnes, 2001). Our present findings suggest that many children may indeed be curious about diverse scientific topics at home, of which many topics seem highly relevant to science content taught in school. Thus, we suggest that teachers should adopt a more inclusive

approach towards their teaching by offering children opportunities in class to exhibit and connect their cognitive and epistemological curiosities to the school curriculum. In addition, we believe that parents may be involved as well to further stimulate and facilitate this transfer. Naturally, teachers would need to learn how to effectively scaffold and guide such conversations with children and their parents.

In sum, while inquiry-based learning approaches have proven to be effective in improving children's question-asking skills, we do not believe that children will develop a genuine understanding of curiosity by simple engagement in hands-on inquiry. As others have stressed as well (e.g., Hodson, 2014), failing to distinguish between these two learning goals may likely lead to confused learning outcomes and misguide teachers' design of curiosity-eliciting lesson activities. Thus, if we want to be successful in addressing the longstanding and valued educational goals to help children become curiosity-driven thinkers (e.g., Lucas et al., 2013; Osborne & Dillon, 2008), we advocate an *explicit* approach to fostering children's curiosity in school, to be regarded as an integral element of the school curriculum. In our view, such an approach would enable teachers to deliberately cultivate a positive attitude towards curiosity and a positive classroom climate, in which children value the epistemic importance of asking questions, derive pleasure from expressing their wonderments and questions, and feel that their epistemic questions are appreciated by their teachers and peers.

Conclusion and future directions

In review, the structured interview procedure we used in the present study proved useful to evaluate children's conceptions and experiences of curiosity. The individual, face-to-face meetings with the children allowed us to explore their responses, provide encouragement and ask for clarifications if needed. Nearly all children provided comprehensive answers to each of our interview questions, including most of the younger children. In addition, our pre-structured interview format also assisted us in keeping the time duration of each interview session less than 10 minutes, which made it possible to conduct the interviews at the schools without disrupting on-going lesson activities too much.

Despite these advantages, we are well aware of the fact that the results presented in this paper are based on data from only two Dutch primary schools. More work is needed to find out how well these findings apply to a wider range of schools. In addition, we should investigate to what extent classroom culture is related to

children's inquiry behavior. Therefore, the development of survey instruments is needed and classroom observation procedures will be useful to explore potential individual differences in the responses of children to curiosity-eliciting lesson activities and guidance by their teachers. Since our study was the first to investigate children's curiosity through the use of children's self-reports, our interview instrument and procedure may also require further validation and optimization in future studies.

Based on the results of the current study, we propose that fostering children's images, beliefs, and feelings about epistemic curiosity in the context of their formal education may prove to be a vital stepping-stone to evoke children's inquiry behavior in the classroom. In fact, based on children's reports in our current sample, we believe that any other efforts to foster children's curiosity in school may turn out to be fruitless, as children seem to generally hold naïve and narrow notions about the nature and educational value of curiosity. In other words, no matter how much we try to stimulate children's curiosity with attractive and diverse curiosity-eliciting lesson activities, if we do not make children aware of *why* we prompt them to be curious, they might just consider these attempts as simple and discrete stimulation strategies and remain reluctant to curiously engage in other learning. The scientific investigation of children's attitudes toward epistemic curiosity may thus allow us to develop measurement instruments that are in line with children's own developing conceptions of curiosity and thereby enable us to design, evaluate and optimize learning activities that aim to foster children's curious thinking in school.

Chapter 3

3 Development and validation of a questionnaire to measure primary school children's images of and attitudes towards curiosity (the CIAC questionnaire)

Abstract

This paper presents a validation study of a questionnaire to measure primary children's images of and attitudes towards curiosity (the CIAC questionnaire). Policy documents and scientific studies on 21st century learning increasingly promote the value of stimulating children's curiosity in primary school. However, no well-established measurement instruments yet exist to assess children's curiosity within educational settings. To fill this void, we focused on the measurement of children's *perceptions* of curiosity, as important precursors to children's potential curiosity-driven behavior. Based on attitude and curiosity theory, we developed seven components of children's images of and attitudes towards curiosity. We translated these components into corresponding measurement scales, which comprise the CIAC. Results of a validation study among 737 children (ages 8-13), using factor analyses, largely confirmed the factor structures of the image and attitude scales and indicated good convergent and discriminant validity. In addition, we provide evidence for the predictive power of children's images and attitudes on their motivation to be curious.

Introduction

This paper presents a validation study of a questionnaire to measure primary children's images of and attitudes towards curiosity (the CIAC questionnaire). Curiosity may be defined as a desire to seek and acquire new information (Berlyne, 1954; Kashdan, 2004; Litman, 2008; Loewenstein, 1994). Historically, Berlyne (1954) and Piaget (1952) were among the first to propose that curiosity may motivate complex exploratory learning behavior. Subsequent research by, for example, Loewenstein (1994), Litman and colleagues (e.g., Litman, Hutchins, & Russon, 2005), and Kashdan and colleagues (e.g., Kashdan & Steger, 2007) has added to the work of Berlyne and Piaget by further defining the dimensionality, determinants, and measures of curiosity.

Recently, research has shifted focus to the investigation of curiosity within educational settings (for reviews on the topic, please see Grossnickle, 2016; Jirout & Klahr, 2012). Policy documents on 21st century learning increasingly promote school curricula that aim to engage children in the scientific process of knowledge development (Lucas, Claxton, & Spencer, 2013; OECD, 2015; Pellegrino & Hilton, 2012). In school settings, children's curiosity is linked to wonderment (e.g., Opdal, 2001), question-asking (e.g., Jirout, 2011), and explanation-seeking behavior (e.g., Litman, Hutchins, & Russon, 2005) and predominantly understood in terms of *epistemic* curiosity: the desire to seek and acquire new intellectual information (Litman & Spielberger, 2003; Loewenstein, 1994; Piotrowski, Litman, & Valkenburg, 2014). Epistemic curiosity is believed to improve children's undertaking of complex inquiry activities (e.g., Von Stumm, Hell, & Chamorro-Premuzic, 2011), their persistence with learning (e.g., Metz, 2008), and their memorization of new information (e.g., Jepma, Verdonschot, Van Steenbergen, Rombouts, & Nieuwenhuis, 2012).

The clear educational value of epistemic curiosity has led many researchers and education policy-makers to advocate the implementation of curiosity-focused pedagogy in primary schools (Claxton & Carr, 2004; Engel, 2006, 2011; Jirout & Klahr, 2012; OECD, 2015; Osborne & Dillon, 2008; Lucas et al., 2013). Unfortunately, however, no well-established guidelines or instruments yet exist to promote or assess children's epistemic curiosity in school settings (for a review on this topic, please see Jirout & Klahr, 2012). Curiosity research has been mostly limited to measuring children's curiosity behavior in response to curiosity-eliciting stimuli (e.g., toys, games) in laboratory settings, rather than in everyday classroom practice. In addition,

proposed curiosity definitions often confound the concept of curiosity with the concepts of interest, intelligence, or motivation, which makes it unclear what is measured or what mechanisms may underlie children's curiosity behavior (Grossnickle, 2016; Silvia & Sanders, 2010).

Studies concerning children's curiosity in primary school settings suggest that the educational content and pedagogy offer children little encouragement to be curious, even as part of inquiry-oriented educational activities (Engel, 2006, 2011; Engel & Randall, 2009; Fortus, 2014). Lesson activities usually emphasize to children the notion that there is just one correct answer to or solution for questions and assignments, and that diverse question-asking and explanation-seeking is disruptive to teachers' pre-scripted instruction (Post & Walma van der Molen, in press; Van Booven, 2015).

It is our belief that children's epistemic curiosity in school can only flourish in a positive classroom climate in which children are taught the epistemic value of being curious-minded, in which they derive pleasure from expressing questions and ideas, and feel that such questions and ideas are appreciated by their teachers and peers. Such perceptions of curiosity can be understood in terms of *attitudes*. Decades of social psychological research show that attitudinal beliefs and affects are important precursors to behavior (e.g., Ajzen, 2001; Ajzen & Fishbein, 1980; Eagly & Chaiken, 1993). In addition, recent research on achievement motivation shows that students' motivational beliefs determine the type of learning strategies they employ (Muis et al., 2015; Muis, Psaradellis, Lajoie, Di Leo, & Chevrier, 2015). According to the control-value theory as proposed by Pekrun (2006), the types of emotions students may experience in school settings depend on their perceptions of control and value evaluations, which can be understood as the degree to which students subjectively *attribute* importance to achievement-related activities (Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011). Thus, on the basis of attitude and motivation theory, we expect that children's images of 'epistemic curiosity' and their perceptions of the value of being curious in school would precede their curiosity behavior in the classroom. Rather than focusing on children's curiosity *behavior*, *skills*, *traits*, or *states*, in this paper, we therefore aimed to investigate children's *perceptions* of curiosity: their images of and attitudes towards curiosity. Shifting the focus towards the investigation of children's attitudes towards curiosity also has some practical advantages. First, attitudes can be measured (Blalock et al., 2008; Reid, 2006). Second, while attitude is considered a relatively stable psychological construct, attitudes can be improved over time (Vogel & Wänke, 2016).

Present study

The current validation study occurred in the context of a larger investigation of the effects of an attitude-focused approach to fostering children's epistemic curiosity in primary schools. For this investigation, we conducted a large-scale intervention study in The Netherlands in which six primary schools participated in a school-wide, curiosity-focused teacher-training program. In addition to measuring changes in teachers' everyday practice over time, the study investigates children's images of and attitudes towards curiosity from the 4th to the 6th grade. To measure changes over time, we developed the Children's Images of and Attitudes towards Curiosity (CIAC) questionnaire. The CIAC provides a comprehensive ensemble of image and attitude components that, according to attitude theory (in particular, the Theory of Planned Behavior; Ajzen, 2001), may precede children's curious behavior in the classroom.

In this paper, we describe the development of the CIAC and present the results of our validation study. Our study included qualitative as well as extensive quantitative methods to determine the convergent and discriminant validity of the developed measurement scales. In the next section, we first outline the theoretical framework that underlies the scales of the CIAC.

Theoretical framework for the CIAC questionnaire

While many different definitions of attitude exist in the literature, the concept of attitude is traditionally described as the psychological 'tendency' of a person to evaluate a particular 'attitude object' in terms of favorable or unfavorable perceptions (Ajzen, 2001; Ajzen & Fishbein, 1980; Eagly & Chaiken, 1993). In this paper, we investigated *epistemic curiosity* as an object of children's attitude in school settings. One of the most well-known models of attitude is the Theory of Planned Behavior (TPB; Ajzen, 1991, 2001; Armitage & Conner, 2001; Godin & Kok, 1996; Hausenblas, Carron, & Mack, 1997).

The TPB generally distinguishes three dimensions of attitude. The first dimension, *Perceptions of Behavioral Attributes,* represents the beliefs and feelings that a person attributes to the attitude object (e.g., the belief that epistemic curiosity fosters one's learning or the pleasurable feeling of posing epistemic questions or ideas in class). The second dimension, *Perceptions of the Social Norm,* describes a person's perception of the social acceptability of the behavior. This may include both a person's negative judgment of others and a person's fear of other people's judgments (e.g., children may fear negative judgments about their own curiosity behavior or they may

judge others for the same behavior). The third dimension, *Perceptions of Behavioral Control,* represents the perceived level of control that a person experiences when performing a certain behavior with respect to the attitude object. In attitude literature, self-efficacy is often regarded as a central component of the *Perceptions of Behavioral Control* dimension of attitude (Ajzen, 1991; Eagly & Chaiken, 1993; Olson & Zanna, 1993). Self-efficacy is a person's perceived capability to perform a behavior when opportunity is provided to do so (e.g., a child's feeling of self-efficacy to pose epistemic questions) (Bandura, 1997).

Together, cognitive, affective, normative, and perceived control perceptions may determine the formulation of a behavioral *intention* to perform or not perform related behavior (e.g., to pose epistemic questions, seek alternative explanations). According to the Expectancy-Value Model (Fishbein & Ajzen, 1974), the strength of this behavioral intention depends on the strength of a person's attitudinal perceptions. In the present study, we asked what cognitive, affective, normative, and control perceptions may constitute important components of children's attitudes towards epistemic curiosity in school settings. Based on a review of research on attitude (e.g., Ajzen, 2001; Eagly & Chaiken, 1993), curiosity (e.g., Engel, 2009, 2011; Grossnickle, 2016; Jirout & Klahr, 2012), scientific literacy (e.g., Osborne & Dillon, 2008; National Research Council, 2012), and lifelong learning and creativity (e.g., Claxton & Carr, 2004; Lucas et al., 2013), we developed a theoretical framework that describes the components of children's attitudes towards epistemic curiosity that seem most relevant for primary school children's curiosity-related behavior.

In addition, we added a separate dimension to the framework that represents children's *images* of curiosity: their mental representation of the term 'curiosity' (Eagly & Chaiken, 1993). These images constitute the 'object' of children's attitudinal evaluation and, therefore, function as an essential determinant of their attitude towards curiosity (Ajzen, 2001; Eagly & Chaiken, 1993; Osborne, Simon, & Collins, 2003). Previous research has shown that many primary school children predominantly associate curiosity with social behavior, such as spying on or prying about others (see Litman & Pezzo, 2007), as opposed to academic learning, such as exploring new subject matter or considering alternative explanations to intellectual problems (Grossnickle, 2016; Post & Walma van der Molen, 2018a). Such a narrow image of curiosity may prevent children from perceiving the epistemic value of curiosity for academic learning. To measure children's images of curiosity and to examine the ways in which children's images are related to their attitudes towards

epistemic curiosity, we included children's social and epistemic images of curiosity as two separate image components in the framework.

Children's images of curiosity

Epistemic image of curiosity. Epistemic images of curiosity portray behaviors of seeking or obtaining new intellectual information (Litman, 2008; Litman et al., 2010; Piotrowski et al., 2014). Examples of such behavior include wanting to know how the human body works or how computers work. In addition, epistemic curiosity could also involve the desire to learn about the epistemology of certain ideas or inventions (e.g., wanting to know how computers were invented).

Social images of curiosity. Social images of curiosity represent behaviors of seeking or obtaining new information about social experiences (Litman & Pezzo, 2007; Litman, Robinson, & Demetre, 2016). Examples of such behavior include spying on or prying about other people. Although children's social curiosity does not directly serve their academic learning, social curiosity is considered to play an important role in children's social development (Grossnickle, 2016), because information about others may help to form friendships (Rosnow, 2001) or to avoid negative social confrontations (Galen & Underwood, 1997).

Children's attitudes towards epistemic curiosity

Personal relevance. We derived the *Personal Relevance* component on the basis of the *Perceptions of Behavioral Attributes* dimension of the TPB. This cognitive component concerns a child's perception of the value of expressing epistemic questions and ideas in class to improve one's own learning. Children will be more likely to engage in curious thinking in school when they perceive the positive outcomes of doing so for their own learning performance (Claxton, 2007; Claxton & Carr, 2004; Lucas et al. 2013; Pellegrino & Hilton, 2012). Thus, we reasoned that children's perceived personal relevance of being curious might constitute an important consideration in whether children do or do not engage in curious thinking in the classroom.

Personal enjoyment. The *Personal Enjoyment* component refers to the pleasurable feeling of asking questions or coming up with new ideas (Arango-Muñoz, 2014; Kashdan & Steger, 2007; Piotrowski et al., 2014). We also derived this affective component of attitude based on the *Perceptions of Behavioral Attributes* dimension of the TPB. It should be noted that the mere pleasure of asking questions opposes typical deficit-type descriptions that define curiosity as a generally unpleasant feeling

of 'not-knowing' that needs to be reduced by learning (for a literature review of opposing views on this topic, please see Grossnickle, 2016). Yet, do we really want children to develop the type of curiosity in school that drives them to reduce their unpleasant feeling of 'not-knowing'? We do not think so. Moreover, while curiosity and enjoyment are conceptualized in the literature as distinct psychological constructs, they do share important features and are believed to both determine children's intrinsic motivation to learn (Reeve, 1989; Grossnickle, 2016). Thus, if we want children to develop an openness to learning, we should show them the joy of question-asking and explanation-seeking, even if their epistemic questions or ideas do not lead to instant answers or solutions. In our view, such an interest-type quality of epistemic curiosity fits better with the 21st century education standards that policy-makers aim to achieve (e.g., Lucas et al., 2013; National Research Council, 2012; OECD, 2015).

Societal relevance. The *Societal Relevance* component was also derived on the basis of the *Perceptions of Behavioral Attributes* dimensions of the TPB. This cognitive component concerns children's perception of the value of curious thinkers to society. Studies have stressed that typical teacher-directed and scripted forms of education may unintentionally convey to children that our collective understanding of the world is already absolute and complete, without revealing to children the tentative nature of such knowledge and the epistemological value of curious thinkers to society (Abd-El-Khalick, 2012; Fouad, Masters, & Akerson, 2015; Trevors, Muis, Pekrun, Sinatra, & Muijselaar, 2017). Such everyday practice may lead children to develop misconceptions about the societal relevance of curiosity (Post & Walma van der Molen, in press). In our view, such misconceptions may prevent them from perceiving the need for curiosity-driven thinkers in society and, thereby, from acting upon their curiosity in school.

Fear of negative judgment. The *Fear of Negative Judgment* component concerns children's fears of their peers' or teachers' negative judgments about being curious in class. We derived the *Fear of Negative Judgment* component on the basis of the *Perceptions of the Social Norm* dimension of the TPB. Research suggests that, in many countries, primary teachers generally feel uncomfortable when children ask diverse questions about topics that teachers themselves do not know the answers to (Ramey-Gassert, Shroyer, & Staver, 1996; Schoon & Boone, 1998; van Aalderen-Smeets, & Walma van der Molen, 2015; Van Booven, 2015). Consequently, many teachers tend to shy away from stimulating children's curious thinking or to judge children's curious questions and ideas as disruptive to their instruction (Claxton &

Carr, 2004; Claxton, 2007). It seems plausible to assume that, over time, this may lead children to develop negative perceptions of what their peers and teachers may think about their curious behavior, which will discourage them from enacting such behavior in the classroom (Marx & Harris, 2006; McCombs, Daniels, & Perry, 2008; Post & Walma van der Molen, in press). If children's implicit negative perceptions are not explicitly attended to by teachers, these may well persist throughout (and beyond) primary school even when teachers provide children opportunity to inquire and curiously explore study subject matter (Post & Walma van der Molen, in press).

Negative opinion. Apart from fearing the judgments by others, we expect that many children may also hold negative judgments about other people's curious question-asking and explanation-seeking behavior. Thus, the *Negative Opinion* component refers to a child's negative opinion about other people's curiosity-related behavior. Similar to the *Fear of Negative Judgment* component, we derived the *Negative Opinion* component on the basis of the *Perceptions of the Social Norm* dimension of the TPB, and likewise expect that children's negative opinions about curious thinkers prevent them from expressing their own epistemic questions and ideas in class.

Self-efficacy. Lastly, the *Self-Efficacy* component refers to children's perceived capability to express epistemic questions or ideas in class when sufficient opportunity is provided. Self-efficacy forms an essential component of the *Perception of Behavioral Control* dimension of the TPB (Ajzen, 1991; Eagly & Chaiken, 1993; Olson & Zanna, 1993) and is widely regarded to be one of the most important determinants of behavior (Bandura, 1997; Palmer, 2006). In line with attitude and self-efficacy theories, we reason that children who perceive themselves to be capable question-askers and explanation-seekers are more likely to carry out such behaviors in comparison to children who feel insecure about their capabilities in this respect.

Hypotheses

We believe that the above-described image and attitude components may represent children's primary considerations to perform (or not perform) epistemic curiosity-related behaviors in school. We differentiate between children's social and epistemic *images* of curiosity in the 'image scale' of our survey. In addition, we propose six components of children's *attitudes* towards epistemic curiosity in our 'attitude scale'. Four attitude components represent perceptions that may positively contribute to children's curiosity behavior (*Personal Relevance*, *Personal Enjoyment*, *Societal*

Relevance, and *Self-Efficacy)*. The other two components represent perceptions that may negatively influence children's curiosity behavior (*Fear of Negative Judgment* and *Negative Opinion)*. We hypothesized that our proposed 'positive components' would correlate positively. Similarly, we expected that our 'negative components' would show a positive correlation. In addition, we expected that the 'positive' and 'negative' attitude components would be either unrelated or negatively correlated.

We also hypothesized, on the basis of attitude theory, that children's image scores would predict their attitude scores (Ajzen, 2001; Eagly & Chaiken, 1993). In particular, we expected that scores on the *Epistemic Image* component would positively predict scores on the *Personal Relevance, Personal Enjoyment, Societal Relevance*, and *Self-efficacy* components, because these attitude components positively relate to the epistemic value of curiosity. We expected that scores on the *Epistemic Image* component would negatively predict scores on the *Negative Opinion* component, because children who associate curiosity with its epistemic use would likely not perceive others' epistemic curiosity behavior to be inappropriate. We expected the *Epistemic Image* component to be unrelated or negatively related to the *Fear of Negative Judgment* component, because children's own epistemic images of curiosity may exist independently from their perceptions of other people's negative opinions about epistemic curiosity behavior. In addition, we expected the *Social Image* component to be unrelated to our attitude components, because children's association of curiosity with social behavior has little relation to the use of curiosity for academic learning and may thus function independently of a positive attitude towards epistemic curiosity.

In line with the TPB, we also considered the predictive power of children's attitude scores on their motivation to be curious (Conner & Armitage, 1998; Fishbein & Ajzen, 1974). Research on curiosity suggests that interest-type curiosity may be positively related to mastery orientation motivation in particular (Grossnickle, 2016; Litman, 2008). However, because our proposed components of children's attitudes towards epistemic curiosity stem from different underlying attitudinal dimensions (i.e., *Perception of Behavioral Attributes*, *Perception of the Social Norm*, and *Perception of Behavioral Control*), we hypothesized that our attitude components might predict different types of motivations. Therefore, we examined the predictive power of each individual attitude component on two largely distinct but relevant motivational components of children's motivation to be curious: children's *Mastery Orientation Motivation* and their *Performance Avoidance Motivation. Mastery Orientation Motivation* is the desire to achieve competence or an understanding, for

the joy and personal use of mastering new tasks (Dweck & Leggett, 1988). *Performance Avoidance Motivation* is the desire to avoid performing, for not revealing one's possible incompetence to others (Darnon, Harackiewicz, Butera, Mugny, & Quiamzade, 2007; Elliot, 1999).

On the basis of the above theories on attitude, curiosity and motivation, we expected that the attitude components that relate to positive perceptions about the personal value of curiosity would positively influence mastery orientation motivation. Thus, we expected scores on the *Personal Relevance, Personal Enjoyment,* and *Self-Efficacy* components to positively predict children's scores on *Mastery Orientation Motivation.* Because our *Societal Relevance* component refers to children's perception about the relevance of 'curious people' to society in general, and not to the relevance of children's *own* curiosity, we expected that this attitude component would be either unrelated or positively related to *Mastery Orientation Motivation.* In addition, we expected our *Fear of Negative Judgment* and *Negative Opinion* components to be either unrelated or negatively related to *Mastery Orientation Motivation,* because children with such negative perceptions about epistemic curiosity will likely feel less inclined to show such behavior themselves.

We expected that scores on our *Fear of Negative Judgment* and *Negative Opinion* components would predict children's performance avoidance motivation, because we hypothesized that children who perceive that (their) epistemic curiosity is inappropriate would avoid expressing their own epistemic curiosity. We expected scores on our *Personal Relevance, Personal Enjoyment,* and *Self-Efficacy* components to be negatively related to *Performance Avoidance Motivation,* because children with positive perceptions about the use of epistemic curiosity would probably actively seek more opportunities to engage in curiosity behavior, rather than avoid them. Because the *Societal Relevance* component does not refer to children's own curiosity, we expected this component to be either unrelated or negatively related to *Performance Avoidance Motivation.*

Finally, some studies suggest that teacher-directed and standardized approaches to learning are limiting children's natural tendency to be inquisitive learners (Claxton & Carr, 2004; Engel, 2006; Engel & Randall, 2009). As we described in the introduction of this paper, such approaches may easily teach children that there is just one correct answer to teachers' questions and that being inquisitive is disruptive to classroom instruction (e.g., Post & Walma van der Molen, in press). This effect might be especially prominent among children who transition towards the upper grades of primary school due to teachers' increasing efforts to help children

pass national high-stake tests by means of 'teaching to these tests' (Jones, Jones, & Hargrove, 2003). Such classroom practice is generally regarded to leave children with little opportunity to curiously explore new learning content on their own and may lead them to develop negative notions about the educational value of curiosity-driven learning (Engel, 2006, 2015; Post & Walma van der Molen, in press). Although little empirical evidence exists to support this proposition, a prevailing hypothesis is that children's attitudes towards epistemic curiosity worsen as they progress through primary school (Engel, 2015). We sought to test this hypothesis in our current study.

Development of the CIAC questionnaire

We followed the framework for construct validity described by Trochim and Donnelly (2006) to develop our questionnaire. Below, we describe in detail what methods we employed to attain and examine construct validity of the CIAC questionnaire.

Establishing translation validity

Content validity. The CIAC questionnaire consists of two separate scales that are based on the image and attitude components described in our theoretical framework: the *Images of Curiosity* scale and the *Attitudes towards Epistemic Curiosity* scale. For each sub-component, we constructed a minimum of four items to allow the removal of possible problematic items later on in the validation process. The original version of the CIAC questionnaire consisted of 41 items.

The *Images of Curiosity* scale includes two hypothesized subscales. The *Social Image* of curiosity subscale aims to measure the extent to which children associate curiosity with questions about social matters, such as wanting to know what you will receive as a birthday present or finding out about other people's personal secrets. The *Epistemic Image* of curiosity subscale was designed to measure to what degree children associate curiosity with cognitive or epistemic questions, such as wanting to know how a computer works or how mathematics was invented. Every question in the *Images of Curiosity* scale has a similar format, stating: 'Suppose you wanted to know how someone obtained knowledge about certain gossip/how the human body works/and so on, indicate how much this has to do with curiosity.' Children's answers were measured on a four-point Likert scale, with response scale options: (1) a very small amount, (2) a small amount, (3) a large amount, (4) a very large amount.

The *Attitudes towards Epistemic Curiosity* scale included six subscales. The first subscale, *Personal Relevance,* was designed to measure to what extent children

perceive the relevance or value of posing epistemic questions at school for their own educational development (e.g., whether they find it important for their own learning to pose epistemic questions). The second subscale, *Personal Enjoyment,* was designed to measure to what degree children enjoy posing epistemic questions about lesson content. The *Societal Relevance* subscale was constructed to measure to what extent children attach societal relevance to epistemic curiosity (e.g., whether curious thinkers foster economic or societal welfare). The fourth subscale, *Negative Opinion*, was constructed to measure whether children hold negative feelings or opinions about other people's epistemic questions (e.g., whether they find curious people to be 'meddlesome' or 'acting smart'). The *Fear of Negative Judgment* subscale was designed to measure the extent to which children fear the negative judgments of others in class (teachers and peers) when expressing epistemic questions or ideas (e.g., the fear that classmates may find you 'nerdy'). The sixth subscale, *Self-Efficacy*, was designed to measure the extent to which children perceive themselves capable of posing epistemic questions about lesson content. Each item in the *Attitudes towards Epistemic Curiosity* scale was formulated as a statement. Children's answers were assessed on a four-point Likert scale, with response scale options: (1) strongly disagree, (2) disagree, (3) agree, (4) strongly agree.

Additional considerations concerning item development. In addition to the specific considerations for item design per subscale, we ensured the clear wording of items and a random sequencing of items in the final format of the questionnaire (Schwarz, 2008). In addition, we linked each item statement to primary school children's everyday context. For example, we used children's own reported experiences with 'being curious' about birthday presents or about things happening in class, which were based on a previous study (Post & Walma van der Molen, 2017). We also ensured that each statement contained a singular, unambiguous, and appropriate attitude object (Ajzen & Fishbein, 2005). For instance, asking children whether they enjoy posing *questions* involves a different 'object' of attitude than asking about their enjoyment of posing *questions to explore new lesson content*.

Likert scale. Although the wording and meaning of the response options differed between the Image and Attitude scales, a four-point Likert response scale was used for both parts of the questionnaire. Although there are different methods for measuring attitudes, such as semantic differential scales, direct interviews, or implicit testing, we chose a Likert scale because this method has several advantages. Likert-scale instruments are suitable when including a large number of items organized in multiple subscales, they can be administered to a large number of respondents, the

items can be answered quickly, and Likert items can be combined to form composite scales that enable parametric testing (Boone & Boone, 2012; Schwarz, 2008).

We used four-point Likert scales, rather than using an uneven number of response options, because primary school children may misinterpret the midpoint in an uneven scale as a *neutral*, *uncertain*, or *do not know* response, which conceptually differs from a midpoint on a sliding scale (Krosnick & Fabrigar, 1997; Kulas & Stachowski, 2009). We kept the number of response options to a minimum, since primary school children often hold fairly dichotomous opinions about all sorts of subject-matter and thus may experience difficulty in responding to scales with too fine-grained response options (Mellor & More, 2014). Four-point Likert response scales have been used effectively in other studies that investigated children's images and attitudes (for examples of such questionnaires, please see Frantom, Green, & Hoffman, 2002; Block, 1995; Post & Walma van der Molen, 2014).

Face validity. After we formulated the items for each hypothesized subscale, we asked several experts in the fields of questionnaire design and attitude research to indicate whether the items were clearly formulated and representative of their corresponding subscales. In addition, several teachers and children from two primary schools that did not take part in the validation study were asked to assess the comprehensibility, clarity, and appropriateness of the items for our target group. Based on the comments we received from the experts, teachers, and children, we identified minor issues with some items and made necessary improvements. Overall, this qualitative pilot study aided us in improving the face validity of the CIAC.

Validation study for the CIAC questionnaire

Respondents

We administered the CIAC questionnaire among a large sample of children ($N = 737$) in the 4th, 5th and 6th grades of nine public primary schools in The Netherlands. The schools were located in different school districts, but were selected on the basis of similar SES background characteristics. The effective sample size included 369 boys (50.1%) and 367 girls (49.9%). Respondents' ages ranged from 8 to 13 years ($M = 10.61$; $SD = .99$).

Six of the nine schools participated in the study as part of a larger school development project that concerned a six-month, school-wide teacher-training program on the topic of inquiry-based teaching. The questionnaire was administered around the time the program had started. The remaining three schools did not

participate in the teacher-training program, but took part in the study on a voluntary basis.

Procedure and materials

A paper-and-pencil version of the questionnaire was administered by a research assistant to all the children in their own classroom during normal school hours. After an introduction for the whole class by the assistant, the children were given the time needed, about 10 minutes, to complete the questionnaire. If a child did not understand a particular item, the researcher provided feedback individually. The objectives, time requirements, and nature of the data collection procedure were explained to the school management, teachers, and parents a few weeks before the start of the data collection. Informed consent was obtained from the parents in accordance with the ethical guidelines of our university.

Data analysis

The criterion validity of the two scales of the CIAC was investigated in several consecutive steps. First, we checked for missing data. We also investigated the discriminant power of each item by evaluating the range of responses and the standard deviation of respondents' scores on each item. Next, using Mplus version 7.4 (Muthén & Muthén, 1998–2015), we performed an exploratory factor analysis (EFA) on the raw (ordinal) data for both the image and attitude scales, using the option 'Categorical' and weighted least squares estimation with means and variance adjusted (WLSMV) and Geomin oblique rotation. Since the CIAC questionnaire had been newly developed on the basis of a new theoretical framework, we first explored and identified the latent factors underlying children's images of and attitudes towards curiosity by using EFA (Prudon, 2015). This approach also helped us to remove any problematic items from the questionnaire before conducting a confirmatory factor analysis (CFA). Unlike EFA, CFA allows for testing model fit, that is, how well the observed data fit a pre-defined hypothesized factor structure (for an extensive review of best practices using both EFA and CFA for instrument validation, please see Prudon, 2015; Schmitt, 2011; Worthington & Whittaker, 2006). CFA was also performed using Mplus and WLSMV estimation.

While Mplus does provide model fit estimations, it does not provide direct information about the convergent and discriminant validity of the subscales under investigation (Carter, 2016). Therefore, we followed the computational formulas provided by Fornell and Larcker (1981) and Raykov (1997) to calculate the necessary

additional measures of Average Variance Extracted (AVE), Composite Reliability (CR), Average Shared Variance (ASV), and Maximum Shared Variance (MSV). Convergent validity of a subscale is considered satisfactory when AVE is equal to or greater than .50 (i.e., the amount of shared variance among items that belong to a subscale). In addition, the CR value of a subscale should be equal to or greater than .70 and greater than the AVE value of the subscale (Fornell & Larcker, 1981). Discriminant validity of a subscale is met when the AVE of the subscale is greater than its MSV and greater than its ASV.

The EFA and CFA analyses were conducted with two different subsamples of the data. A random sampling procedure was used to extract subsample 1 ($n = 368$) and subsample 2 ($n = 369$) from the total respondent group. Equivalence of the subsamples was investigated for gender and grade using chi-square tests. Results indicated that boys and girls ($\chi^2 = .60$, *df* = 1, *p* = .46) and children from different grades ($\chi^2 = 2.20$, *df* = 2, *p* = .33) were equally distributed across both subsamples.

As will be described in the Results section, we also assessed the ability of the CIAC questionnaire to make fair comparisons between the image and attitude scores of children from different grade levels, since we assume that the CIAC questionnaire measures the same constructs the same way for children across different grade levels. To test this assumption, we performed a confirmatory factor analytic test of measurement invariance (Vandenberg & Lance, 2000).

Finally, we assessed the predictive validity of the CIAC by examining the extent to which children's scores on the *Epistemic Image* subscale predicted their scores on each separate attitude component and the extent to which children's attitudes scores predicted their scores on *Mastery Orientation Motivation* and *Performance Avoidance Motivation.* To this end, we used the validated Achievement Goal Questionnaire by Elliot and McGregor (2001) to assess children's goal achievement motivations (see also Pekrun et al., 2011). Predictive validity was examined by use of Structural Equation Modelling (SEM) through fitting our hypothesized models of relationships among the image, attitude and motivation variables to our acquired data (Muthén & Muthén, 1998–2015). In addition, using multivariate analysis of variance (MANOVA) in SPSS version 24, we examined children's overall attitude scores and tested whether the children in Grade 6 showed decreased attitude scores in comparison to the children in lower grades.

Results

Preliminary data checks

We carried out preliminary data checks to detect the presence of missing data. The percentage of missing data was 2.77% in subsample 1 and 2.93% in subsample 2. Little's (1988) MCAR test results indicated that the missing data were missing completely at random in subsample 1 (χ^2 = 3394.66; df = 3277; p = .07) and subsample 2 (χ^2 = 2906.911; df = 2901; p = .47). Both EFA and CFA were performed with the raw data from subsample 1 and 2 that included these missing data, using the default procedure for handling missing data in MPlus. We also examined the discriminant power of each item by computing the standard deviation and range of responses. For each item, the standard deviation should hover around 1 and each response option should be used at least once (Coulson, 1992; Schwarz, 2008). The standard deviations of all items in our questionnaire ranged between .77 and .97 and all response options were used at least once, indicating sufficient discriminant power. The data were thus considered adequate for subsequent factor analyses.

Exploratory factor analysis

Iterative exploratory factor analyses with WLSMV were conducted on subsample 1 to investigate the latent factor structure in the data. Since we expected the subscales to correlate, Geomin oblique rotation was used to determine the best rotated solution (Reise, Waller, & Comrey, 2000). No maximum number of factors was preset. Items were omitted from consecutive EFAs if they showed a factor loading lower than .40 or cross loadings less than a .15 difference from an item's greatest factor loading (Floyd & Widaman, 1995; Hair, Black, Babin, Anderson, & Tatham 2006; Worthington & Whittaker, 2006).

Images of curiosity. We first performed EFA on the items of the *Images of Curiosity* scale. Two items that belonged to the *Social Image* subscale were removed from the questionnaire because of poor factor loadings (< .40). Subsequent EFA analysis revealed the presence of two factors with Eigenvalues above 1. As shown in Table 3.1, all items designed to address each of the two image components loaded onto one particular factor, resulting in a two-factor structure that corresponds to the two hypothesized subscales of the questionnaire design. Factor one (Eigenvalue = 2.94) contains the four items that refer to children's *Social Image* of curiosity. Factor two (Eigenvalue = 1.66) includes the five items that refer to children's *Epistemic Image* of

curiosity. The factor loadings ranged between .48 and .76. Please see Table 3.1 for the obtained factor structure and factor loadings of the *Images of Curiosity* scale.

     Attitudes towards epistemic curiosity. We conducted separate exploratory factor analyses for the items on the *Attitudes towards Epistemic Curiosity* scale. After several EFA iterations, three items were removed from the questionnaire because of poor factor loadings or cross loadings, leaving a total of 18 items. The final factor solution consisted of five factors with Eigenvalues above 1. The factor loadings ranged between .50 and .80. However, the observed factor structure differed in two ways from our hypothesized factor structure. Below, we outline these differences and describe what adaptations we made to our subscales on the basis of the EFA.

     First, while we hypothesized that the *Personal Relevance* and *Personal Enjoyment* subscales would exist as independent subscales, the EFA revealed that both constructs loaded on a joint factor. In line with the Theory of Planned Behavior (TPB), this result may not be too surprising, since the TPB conceptualizes both cognitive and affective perceptions as part of an overarching concept of 'attitudinal perceptions of behavioral attributes' (Ajzen, 1991). Unlike the *Societal Relevance* factor – which refers to the value that children may attribute to people's epistemic curiosity for society *in general* – the *Personal Relevance* and *Personal Enjoyment* factors both refer to children's perceptions of the value of epistemic curiosity for their *own* development. Therefore, in accordance with the obtained EFA factor structure, we decided to unify the hypothesized *Personal Relevance* and *Personal Enjoyment* subscales into a revised, joint subscale, which we labeled *Personal Inclination*. The *Personal Inclination* subscale contains seven items that cover possible cognitive and affective aspects of the learning-related value that a child may attach to his or her own epistemic curiosity.

     Second, the EFA revealed that our hypothesized *Fear of Negative Judgment* subscale seemed to be more accurately described in terms of children's *Fear of Classmates' Negative Judgment* (Eigenvalue = 1.05), because the three items of the original subscale that survived the EFA relate particularly to children's fear of their *classmates'* judgments, rather than possible fears of their *teacher's* judgments. The items about teachers' judgments did not load onto a distinct factor of their own. On the basis of the factor structure obtained by the EFA, we derived the following five attitude subscales: *Personal Inclination* (revised) (Eigenvalue = 4.98); *Societal Relevance* (Eigenvalue = 1.42); *Negative Opinion* (Eigenvalue = 1.27); *Fear of Classmates' Negative Judgment* (Eigenvalue = 1.05); and *Self-Efficacy* (Eigenvalue

= 2.62). Please see Table 3.2 for the factor structure and factor loadings of the *Attitudes towards Epistemic Curiosity* scale.

Confirmatory factor analysis

We performed CFA with WLSMV to assess how well the data of subsample 2 fitted the obtained factorial structures for the image and attitude subscales that we derived by EFA. To determine model fit, we examined multiple goodness-of-fit indices. Because the χ^2 statistic is highly sensitive to large respondent groups, we supplemented the conventional model fit indices with the absolute model fit estimate Weighted Root Mean Square Residual (WRMR) (Prudon, 2015; Worthington & Whittaker, 2013; Yu, 2002). The following model fit indices were used to examine whether the factor structures of our scales fitted the data: (1) the WRMR, (2) the Comparative Fit Index (CFI), (3) the Tucker-Lewis Index (TLI), and (4) the RMSEA (Floyd & Widaman, 1995). WRMR should be below 1.0 to indicate good fit, CFI and TLI values should exceed .95, and RMSEA values should be below .08 (Hu & Bentler, 1999; Mueller & Hancock, 2008; Yu, 2002). It should be noted that the above model fit threshold values are simply guidelines and should not be interpreted as strict rules (Prudon, 2015). In addition, we calculated the AVE, CR, MSV and ASV values for each subscale as obtained by CFA to assess their convergent and discriminant power (Carter, 2016).

Images of curiosity. We performed CFA on the basis of the factor structure solution that had been previously obtained by EFA. The CFA results revealed some additional poor factor loadings for items that belonged to the *Social Image* subscale. These items were then removed from the analysis, leaving a total of seven items for the *Images of Curiosity* scale, two items for the *Social Image* subscale and five items for the *Epistemic Image* subscale. Although more than two items should ideally represent a subscale, the use of two items per subscale is considered satisfactory when the items are fairly strongly correlated by measure of Spearman-Brown correlation (Eisinga, Grotenhuis, & Pelzer, 2013; Worthington & Whittaker, 2006). The two items of the *Social Image* subscale are reasonably correlated (r = .56), so we decided to keep this factor with these two items. Subsequent CFA of the revised *Images of Curiosity* scale confirmed good model fit (WRMR = .60; CFI = .99; TLI= .98; RMSEA = .06). Please see Table 3.1 for the factor structure and factor loadings of the *Images of Curiosity* scale as obtained by CFA. Inspection of the factor correlation matrix (see Table 3.3) showed that the *Social Image* and *Epistemic Image* of curiosity subscales were marginally correlated (r = .38).

Table 3.1

Factor structure solution for the Images of Curiosity scale as obtained by exploratory and confirmatory factor analysis

Indicate how much curiosity has to do with…	EFA[a]		CFA[b]	
	Social	Epistemic	Social	Epistemic
wanting to know how someone found out about a secret	.76		.81	
wanting to know how someone found out about a certain rumor	.60		.76	
(D) wanting to figure out what others are thinking or feeling	.50		–	
(D) eavesdropping on private conversations to figure out what they're talking about	.48		–	
wanting to know how the human body works		.73		.66
wanting to know how a car works		.69		.76
wanting to know how computers were invented		.57		.63
wanting to know how birds are able to fly		.56		.58
wanting to do how math was invented		.55		.62

Note:

[a] Exploratory factor analysis (EFA) conducted with weighted least squares estimation with means and variance adjusted and Geomin rotation with subsample 1. Values represent rotated factor loadings.

[b] Confirmatory factor analysis (CFA) conducted with weighted least squares estimation with means and variance adjusted with subsample 2. Values represent factor loadings.

Factor loadings are only displayed for items with loadings > .40 on their expected factors and with cross loadings > .15 difference from the item's greatest factor loading.

(D): indicates that the item was eventually deleted from the questionnaire due to poor factor loading (< .40) based on results obtained by CFA.

Table 3.2

Factor structure solution for the Attitudes towards Epistemic Curiosity scale as obtained by exploratory and confirmatory factor analysis

Item	EFA[a]					CFA[b]				
	PI	SR	NO	FCNJ	SE	PI	SR	NO	FCNJ	SE
I really like to wonder about all the things I learn at school	.72					.67				
It is very important to me to come up with interesting questions at school, because then I learn more about the things around me	.70					.65				
It is very important to me to wonder about lots of things in class, because then I learn more about all sorts of different things.	.59					.70				
I really like to ask questions about all sorts of things in class	.55					.62				
I think people who want to know a lot are very important to the economy of		.77					.67			
I think people who ask good questions have a big impact on society		.66					.65			
I think people who often come up with interesting questions are very important to society		.65					.77			
I feel classmates are being stubborn when they always want to know all about everything in class			.67					.60		
I find classmates to be annoying when they ask a lot of smart questions in class			.65					.64		

Table 3.2 continued

Item	EFA	CFA
I feel people who ask a lot of questions come across as dumb	.60	.69
I'm afraid that my classmates will think I'm a nerd if I ask a lot of smart questions in class	.79	.71
I'm afraid that my classmates will think it's stupid if I want to know more about something we're learning in class	.71	.77
I find it scary to show that I'd like to know more about a topic in class	.58	.69
I am really good at coming up with smart questions in class	.80	.67
I am really good at coming up with new questions about all sorts of topics in lessons at school	.69	.76
I think I am really good at figuring out new things at school	.53	.76
I am really good at coming up with smart questions about all sorts of subjects at school	.50	.65

Note:

[a] Exploratory factor analysis (EFA) conducted with weighted least squares estimation with means and variance adjusted and Geomin rotation with subsample 1. Values represent rotated factor loadings.

[b] Confirmatory factor analysis (CFA) conducted with weighted least squares estimation with means and variance adjusted with subsample 2. Values represent factor loadings.

Factor loadings are only displayed for items with loadings > .40 on their expected factors and with cross loadings > .15 difference from the item's greatest factor loading.

PI = the Personal Inclination subscale, SR = the Societal Relevance subscale, NO = the Negative Opinion subscale, FCNJ = the Fear of Classmates' Negative Judgment subscale, SE = the Self-Efficacy subscale.

65

As depicted in Table 3.3, AVE values for the *Social Image* subscale (AVE = .74) and the *Epistemic Image* subscale (AVE = .55) indicated sufficient convergent validity of the subscales. In addition, CR values for both subscales exceeded the threshold value of .70 and exceeded the AVE values for each respective subscale. The discriminant power of the subscales was also found to be sufficient, as indicated by AVE values that exceeded MSV and ASV.

Attitudes towards epistemic curiosity. We performed CFA on the basis of the factor structure solution that had been previously obtained by EFA. The CFA results confirmed good model fit (WRMR = .84; CFI = .98; TLI = .97; RMSEA = .04). Please see Table 3.2 for the factor structure and factor loadings of the *Attitudes towards Epistemic Curiosity* scale as obtained by CFA. However, our subsequent examination of the convergent and discriminant power of the attitude subscales did reveal two minor issues. As for the convergent power of individual subscales, Table 3.3 shows that while the AVE and CR values of most of the attitude subscales exceeded the threshold values of .50 and .70 respectively (thus indicating that the subscales possess sufficient convergent power), the CR value of the *Negative Opinion* subscale (.68) was slightly lower than .70.

As for the discriminant power of the individual attitude subscales, Table 3.3 shows that the AVE did exceed the ASV, but not the MSV, for the *Personal Inclination* and *Self-Efficacy* subscales. This result indicates that both subscales share considerable variance, which may suggest that the items of the *Personal Inclination* and *Self-Efficacy* subscales were interpreted by children as being conceptually similar. However, because MSV only marginally exceeded AVE and the discriminant power of the subscales was corroborated by adequate convergent power, model fit and conceptual design, we think that the amount of shared variance of the *Personal Inclination* and *Self-Efficacy* subscales is acceptable.

The factor correlation matrix (see Table 3.3) revealed that the subscales correlated as we expected. The positive attitude components, *Personal Inclination*, *Societal Relevance*, and *Self-Efficacy*, all showed statistically significant positive correlations. Similarly, the negative attitude components, *Negative Opinion* and *Fear of Classmates' Negative Judgment*, also showed statistically significant positive inter-correlations. In addition, as hypothesized, the factor correlation matrix shows that positive and negative attitude subscales were either unrelated (insignificant factor correlations) or negatively correlated. These findings further add to the construct validity of the CIAC questionnaire.

Table 3.3
Factor correlations, CR, AVE, MSV, ASV, scale means and standard deviations for the Image and Attitude subscales as obtained by CFA

	Images of Curiosity		Attitudes towards Epistemic Curiosity				
	Social	Epistemic	Personal Inclination	Societal Relevance	Negative Opinion	Fear of Classmates' Negative Judgment	Self-Efficacy
Images of Curiosity							
Social Image	–						
Epistemic Image	.38**	–					
Attitudes towards Epistemic Curiosity							
Personal Inclination			–				
Societal Relevance			.70**	–			
Negative Opinion			-.38**	-.28**	–		
Fear of Classmates' Negative Judgment			-.07	.11	.62**	–	
Self-Efficacy			.77**	.60**	-.17*	-.03	–
Composite Reliability (CR)	.76	.79	.76	.74	.68	.77	.80
Average Variance Extracted (AVE)	.74	.55	.56	.61	.57	.66	.63
Maximum Shared Variance (MSV)	.14	.14	.59	.49	.38	.38	.59
Average Shared Variance (ASV)	.14	.14	.31	.24	.16	.10	.25
Subscale mean	2.98	2.67	2.65	2.71	1.94	1.82	2.46
Subscale standard deviation	.77	.62	.62	.62	.67	.68	.64

Subscale means are weighted averages and scores could range between 1 to 4
* Factor correlation is statistically significant at $p < .05$
** Factor correlation is statistically significant at $p < .01$

Measurement invariance analysis

Next, we conducted multiple-group CFA to test measurement invariance of the CIAC questionnaire across the 4th (n = 214), 5th (n = 242) and 6th (n = 281) grade groups, using the total respondent sample (N = 737). We examined configural invariance (i.e., equality of factor structures) and metric invariance (i.e., equality of factor loadings) of the observed data for both the *Images of Curiosity* and *Attitudes towards Epistemic Curiosity* scales. To assess configural invariance, we allowed item parameters (i.e., factor loadings, item intercepts, and item uniqueness), factor variances, and latent means to vary freely across groups. To assess metric invariance, we fixed the factor loadings of the factor structure across groups.

We examined changes in CFI (ΔCFI) and χ^2 ($\Delta\chi^2$) as our primary tests of measurement invariance, where a ΔCFI less than or equal to .01 indicates invariance and non-significant p-values (< .05) for the $\Delta\chi^2$ between the measurement models indicate invariance (Cheung & Rensvold, 2002). As presented in Tables 3.4 and 3.5, for each of our measurement invariance tests per grade level group comparison, the ΔCFI did not exceed .01 and the p-values for $\Delta\chi^2$ were non-significant. We therefore conclude that both the *Images of Curiosity* and *Attitudes towards Epistemic Curiosity* scales may be used to compare scores across the 4th, 5th, and 6th grade groups.

Predictive validity

Predictive validity of Images on Attitudes. We assessed the ability of the *Epistemic Image* subscale to predict children's attitude scores, as described in our Hypothesis section. Predictive validity was examined by testing the hypothesized structural model using SEM analysis, which indicated good model fit (WRMR = 1.07; CFI = .97; TLI = .96; RMSEA = .04). Table 3.6 summarizes the observed regression coefficients among the image and attitude components. As hypothesized, scores on the *Epistemic Image* component significantly predicted scores on the attitude components Personal Inclination (R^2 = .19, p < .01), *Societal Relevance* (R^2 = .12, p < .01), and *Self-Efficacy* (R^2 = .13, p < .01). In addition, the *Epistemic Image* component negatively predicted the *Negative Opinion* component (R^2 = .02, p = n.s.). The *Epistemic Image* component was unrelated to the *Fear of Classmates' Negative Judgment* component (R^2 = .00, p = n.s.). The *Social Image* component did not predict children's scores on any of the attitude components (R^2 < .04, p > .06), which confirms the hypothesis that children's attitudes towards epistemic curiosity are independent of children's social images of curiosity.

Table 3.4

Measurement invariance results of the Images of Curiosity scale for the total respondent sample, with grade level as the group comparison

Group comparison	χ^2	df	TLI	CFI	RMSEA	ΔCFI	$\Delta\chi^2$	Δdf	p
Grade 4 vs. Grade 5									
Configural invariance	34.835	26	.99	.99	.04	–	–	–	–
Metric invariance	43.947	31	.98	.99	.04	.00	8.702	5	.12
Grade 5 vs. Grade 6									
Configural invariance	45.023	26	.99	.99	.05	–	–	–	–
Metric invariance	47.597	31	.99	.99	.05	.00	4.773	5	.44
Grade 4 vs. Grade 6									
Configural invariance	48.735	26	.98	.99	.06	–	–	–	–
Metric invariance	53.414	31	.98	.99	.05	.00	6.786	5	.24

Table 3.5

Measurement invariance results of the Attitudes towards Epistemic Curiosity scale for the total respondent sample with grade level as the group comparison

Group comparison	χ^2	df	TLI	CFI	RMSEA	ΔCFI	$\Delta\chi^2$	Δdf	p
Grade 4 vs. Grade 5									
Configural invariance	375.959	218	.95	.96	.06	–	–	–	–
Metric invariance	390.616	230	.95	.96	.06	.00	18.735	12	.10
Grade 5 vs. Grade 6									
Configural invariance	394.665	218	.96	.97	.06	–	–	–	–
Metric invariance	409.750	230	.96	.96	.06	.01	18.154	12	.11
Grade 4 vs. Grade 6									
Configural invariance	429.956	218	.94	.95	.06	–	–	–	–
Metric invariance	442.284	230	.94	.95	.06	.00	18.877	12	.09

Table 3.6
Summary of regression analyses for the Image components predicting the Attitude components (N = 737)

	Personal Inclination			Societal Relevance			Negative Opinion			Fear of Classmates' Negative Judgment			Self-Efficacy		
	B	SE B	β	B	SE B	β	B	SE B	β	B	SE B	β	B	SE B	β
Social Image	−.01	.05	−.01	.09	.05	.10	.04	.05	.05	.02	.05	.02	−.01	.05	−.01
Epistemic Image	.46	.05	.43**	.32	.06	.32**	−.15	.05	−.16**	.01	.05	.01	.39	.05	.37**

* Beta is statistically significant at $p < .05$
** Beta is statistically significant at $p < .01$

These findings thus reveal that the *Social Image* and *Epistemic Image* components predicted results that they were hypothesized to predict, thereby supporting the predictive validity of the *Images of Curiosity* scale.

Predictive validity of Attitudes on Motivations. We examined the ability of the *Attitudes towards Epistemic Curiosity* subscales to predict scores on *Mastery Orientation Motivation* and *Performance Avoidance Motivation*. In preparation of this analysis, we first performed CFA to test the supposed two-factor structure underlying the AGQ items from Elliot and McGregor (2001). The CFA results confirmed that the items that belonged to the *Mastery Orientation Motivation* scale and the items that belonged to the *Performance Avoidance Motivation* scale fitted the data well (WRMR = .99; CFI = .99; TLI = .98; RMSEA = .06). Tests of convergent power of each motivation scale proved to be sufficient, as indicated by AVE and CR values that were greater than .50 and .70 respectively. In addition, we assessed the discriminant power of the motivation scales compared to the attitude subscales to exclude the possibility of multicollinearity. The results showed that the amount of shared variance among the separate motivation scales and attitude subscales was acceptable, as indicated by AVE values of each motivation scale that exceeded MSV and ASV. In sum, these preparatory analyses showed that the motivation scales can be regarded as sufficiently independent from each other and from the attitude subscales and that, thereby, these latent variables are suitable to be part of a measurement model to test our hypothesized structural models (Prudon, 2015).

To examine the extent to which scores on the *Attitudes towards Epistemic Curiosity* components predicted children's *Mastery Orientation Motivation* and *Performance Avoidance Motivation*, we tested a unified structural model that related each attitude component to each motivation factor. Table 3.7 summarizes the observed regression coefficients among the attitude components and motivation factors. The observed statistical relationships among the attitude components and the *Mastery Orientation Motivation* factor (R^2 = .50, p < .01) and the relationships among the attitude components and the *Performance Avoidance Motivation* factor (R^2 = .33, p < .01), largely confirmed our hypotheses (WRMR = 1.21; CFI = .96; TLI = .96; RMSEA = .04). Scores on the *Personal Inclination* component significantly predicted scores on *Mastery Orientation Motivation*. In addition, the components *Societal Relevance*, *Negative Opinion*, and *Fear of Classmates' Negative Judgment* related to *Mastery Orientation Motivation* as expected as well. However, in contrast

to our predictions, *the Self-Efficacy* component showed no statistically significant relationship with *Mastery Orientation Motivation.*

As hypothesized as well, the components *Personal Inclination* and *Societal Relevance* showed to be negatively related to *Performance Avoidance Motivation.* In addition, the scores on the *Negative Opinion* and *Fear of Classmates' Negative Judgment* components significantly predicted children's *Performance Avoidance Motivation.* However, contrary to our predictions, the *Self-Efficacy* component showed no statistically significant relationship with *Performance Avoidance Motivation.* In sum, these findings reveal that, for the most part, children's attitudes towards curiosity predict their mastery orientation and performance avoidance motivations and thereby support the predictive validity of the *Attitudes towards Epistemic Curiosity* scale.

Table 3.7
Summary of regression analyses for the Attitude components predicting the Motivation components (N = 737)

	Mastery Orientation Motivation			Performance Avoidance Motivation		
	B	*SE B*	β	*B*	*SE B*	β
Personal Inclination	.81	.12	.73**	−.27	.12	−.24*
Societal Relevance	−.09	.08	−.07	−.18	.08	−.15*
Negative Opinion	−.18	.09	−.13*	.38	.09	.28**
Fear of Classmates' Negative Judgment	−.10	.07	−.09	.22	.07	.19**
Self-Efficacy	−.09	.09	−.07	.09	.09	.07

* Beta is statistically significant at *p* < .05
** Beta is statistically significant at *p* < .01

Children's attitude scores per grade level. Lastly, we investigated the degree to which children's attitude scores declined as a function of children's grade level. To this end, we first computed children's weighted sum scores for each attitude subscale per grade level (see Table 3.8). Notably, these data reveal that, on average, children's scores for the *Personal Inclination*, *Societal Relevance*, and *Self-Efficacy* components lied around the scale's midpoint (2.5), which indicates that children's

attitudes towards these matters in the present sample, irrespective of their grade level, were generally only moderate at best. Contrary, children's scores on *Negative Opinion* and *Fear of Classmates' Negative Judgment* lied below the scale's midpoint, which indicates that, on average, children did not so much perceive their classmates' epistemic curiosity behavior in negative terms, nor did they perceive that their classmates negatively judged their epistemic curiosity behavior.

We further examined differences in children's attitude scores per grade level by performing MANOVA with grade level as the between-subject factor and the five attitude components as multivariate dependent variables. The omnibus test of between-subject effects using Wilk's statistic revealed a significant main effect of grade, $\Lambda = .93$, $F(10, 1436) = 5.385$, $p = .00$, $\eta^2 = .04$. As shown in Table 3.8, post-hoc univariate analyses for each separate attitude component confirmed statistically significant but small differences only between the attitude scores of children from Grade 6 with the children from either Grade 4 or Grade 5 for the attitude components *Personal Inclination*, *Societal Relevance*, and *Self-Efficacy*. In these cases, the attitude scores of children from Grade 6 were somewhat lower than the scores of children from Grade 4 and Grade 5. This result suggests that, on the basis of the present data sample, children's perceptions about the learning value and use of (their) epistemic curiosity may indeed decline as they progress through primary school.

Table 3.8
Comparison of children's weighted sum scores for each separate attitude component per grade level

Attitude subscales	Grade 4 (n = 214)		Grade 5 (n = 242)		Grade 6 (n = 281)		Grade 4 vs Grade 5		Grade 5 vs Grade 6		Grade 4 vs Grade 6	
	M	SD	M	SD	M	SD	η^2	p	η^2	p	η^2	p
Personal Inclination	2.81	.64	2.69	.59	2.51	.62	–	n.s.	.03	.00	.05	.00
Societal Relevance	2.83	.64	2.75	.62	2.52	.65	–	n.s.	.03	.00	.05	.00
Negative Opinion	1.97	.70	1.84	.66	1.93	.67	–	n.s.	–	n.s.	–	n.s.
Fear of Classmates' Negative Judgment	1.83	.76	1.83	.67	1.79	.64	–	n.s.	–	n.s.	–	n.s.
Self-Efficacy	2.57	.68	2.45	.62	2.27	.61	–	n.s.	.02	.01	.05	.00

Note: Weighted sum scores could range between 1 (strongly disagree) to 4 (strongly agree).

Discussion

This study focused on the development and validation of a questionnaire to measure primary school children's images of curiosity and their attitudes towards epistemic curiosity. To the best of our knowledge, the development of the CIAC is the first attempt to measure these aspects in children. Based on curiosity and attitude research, we expect that a fruitful approach to stimulating children's epistemic wonderment, questions, and ideas in the classroom is to foster their positive, epistemic images of curiosity and their positive beliefs and feelings about being curious learners in school. To measure these images and attitudes validly and reliably over time, we developed and tested the CIAC.

Because the CIAC questionnaire and its underlying theoretical components had not been empirically tested before, we employed both qualitative and extensive quantitative methods to verify the construct validity of the CIAC questionnaire. We also assessed the ability of the CIAC questionnaire to measure the same image and attitude components the same way for children across the 4th, 5th and 6th grades (Hirschfeld & Brachel, 2014; Milfont & Fischer, 2010; Vandenberg & Lance, 2000).

Main findings of the study

The results of our study provide good evidence to support the construct validity of the CIAC. Although our results showed some minor deviations from the constructs that we originally hypothesized, overall, our data fitted the main underlying dimensions of the Theory of Planned Behavior well. Furthermore, results showed that the CIAC demonstrated full configural and metric measurement invariance for children across the middle and upper grades of primary school. In addition to the CIAC's internal validity, the observed relationships between the image, attitude, and motivation variables proved to be largely consistent with theory (e.g., Eagly & Chaiken, 1993; Osborne, Simon, & Collins, 2003) and thus provided support for the predictive validity of the CIAC as well. Also, in accordance with theories proposed by others in the field (e.g., Engel, 2006; Jirout & Klahr, 2012), the attitude scores revealed that the children in our sample generally did not hold very positive perceptions about the value and use of being curious about new subject matter in school and that the children in Grade 6 generally felt less positive about these matters than the children in the lower grades. Below, we discuss some issues that might improve certain aspects of the CIAC.

__Images of curiosity.__ While EFA revealed that the items of the *Social Image* subscale all loaded on the specified factor, CFA indicated that two of the four items

loaded poorly (< .40), leaving a total of only two items for the *Social Image* subscale. This finding stresses the importance of employing CFA as well as EFA to thoroughly assess the factor structures obtained by EFA. Research indicates that it is not ideal to represent a subscale by only two items (Worthington & Whittaker, 2006). Thus, the number of items of the *Social Image* subscale should be increased in an improved version of the CIAC, with items that better represent children's social images of curiosity. It is plausible to assume that this addition of items will lead to improved convergent and discriminant power of the *Social Image* subscale. Based on previous research by Litman and Pezzo (2007) on the measurement of social curiosity, examples of such items may be: 'Suppose you wanted to figure out what others are thinking or feeling, indicate how much this has to do with curiosity'; 'Suppose you try to figure out what someone is hiding from you, indicate how much this has to do with curiosity'; 'Suppose you eavesdrop on a private conversation to figure out what is being talked about, indicate how much this has to do with curiosity'.

Attitudes towards epistemic curiosity. Our factor-analytic examination of the *Attitudes towards Epistemic Curiosity* scale indicated that most of our developed items loaded on their expected latent factors. However, EFA did produce some unexpected results. First, the EFA indicated that the items that were originally developed for the *Personal Relevance* and *Personal Enjoyment* subscales loaded on one, joint factor. As we already described in the Results section, this finding is in line with the proposition in the TPB that the attitudinal dimension 'Perceptions of Behavioral Attributes' comprises both cognitive and affective perceptions (Ajzen, 1991). Therefore, we decided to continue our analyses with a revised scale, which we labeled *Personal Inclination*. CFA confirmed that the *Personal Inclination* subscale fitted the data well.

Second, the results of the EFA indicated that *Fear of Negative Judgment* predominantly consisted of children's *Fear of Classmates' Negative Judgment*. In retrospect, it is not surprising that children may attribute different social norms for 'being curious' in class to their peers or to their teachers. This finding again underlines the value of employing both EFA and CFA. CFA indicated that the *Fear of Classmates' Negative Judgment* fitted the data well and that the subscale possessed sufficient convergent and discriminant power.

However, this result does not necessarily imply that children *only* fear their classmates' judgments. To investigate whether an improved version of the CIAC should include a separate scale that measures children's fears of their teachers' judgments, we conducted follow-up interviews with groups of children from the 4th,

5th, and 6th grades from one primary school that participated in the present study. Each group of children consisted of two boys and two girls who were randomly selected from each grade. We first asked the children to what extent they ever experienced having epistemic questions or ideas about lesson topics in class. Second, we asked them if they ever felt afraid of posing their curious questions in class. All children reported having regular curious questions about lesson topics in class. About half of the children in the sample indicated being afraid to express such curious questions. Of these children, all of them indicated that they predominantly feared their peers as likely to make fun of their curiosity. In contrast, the children in this sample did not fear the possibility of their teacher judging their curious questions or ideas in a negative way. Given these additional findings, we propose that children's fear of classmates' negative judgments is indeed an important underlying attitudinal dimension that may hinder children's curiosity behavior, while their fear of teachers' negative judgments seems to be a less important underlying attitudinal factor.

Lastly, the observed relationships among the attitude and motivation variables largely confirmed our hypotheses, with *Self-Efficacy* as the only exception. Children's self-efficacy scores appeared to be unrelated to their mastery orientation motivation and performance avoidance motivation. This result is surprising, because self-efficacy is found to be a central predictor of motivation in a wide range of studies on attitude (Ajzen, 1991; Eagly & Chaiken, 1993; Olson & Zanna, 1993) and achievement (Bandura, 1997; Pekrun, 2006). A tentative explanation could be that the children in our sample had little experience with curiosity-focused or inquiry learning. An explorative study by Post and Walma van der Molen (in press) suggests that children may indeed only barely engage in curiosity-driven learning in primary school. This may have made it difficult for the children to rate their efficacy and might explain why children's self-efficacy scores did not relate to their motivation scores.

Directions for future research

An improved version of the CIAC should be re-validated by administering the questionnaire again among a large group of primary school children. In addition, it should be noted that the current version of the CIAC was developed for children in the 4th, 5th, and 6th grade. We expect that the CIAC is not suited for children younger than 8 or 9 years old, because the survey items might be too demanding for younger children to comprehend. However, we do expect that the CIAC is useful for measuring older children's images and attitudes (up to 14 or perhaps even 15 years old), which would allow researchers and educators to use the CIAC to track children's developing

images and attitudes over longer periods of time. Nonetheless, this expectation needs to be tested as well, by means of factor-analytic tests of measurement invariance that examine the equality of factor structures and factor loadings of children across these different age groups.

The current version of the CIAC consists of Dutch items, which were developed and refined on the basis of the feedback that we received from Dutch primary school children and teachers. Although we have carefully translated the items of the CIAC from Dutch to English, cross-cultural validation studies should determine the construct validity of the CIAC for children in other countries. To this end, we invite fellow researchers to collaborate with us and validate a translated and improved version of the CIAC in their country.

It is also worth noting that the attitude components *Societal Relevance* and *Fear of Classmates' Negative Judgment* both appeared to be unrelated to *Mastery Orientation Motivation*, as we hypothesized. In our view, the absence of these relationships may be explained by the fact that both components refer to children's perceptions about the effects that *other people's* behavior may have, rather than to the effects of children's own behavior. Nevertheless, we do believe that both attitude components constitute important elements of children's attitudes towards curiosity. To assess the predictive power of these attitude components on motivation, however, we suspect that other types of goal motivational measures might be more appropriate, such as performance approach motivation (Reeve, 2015). For example, children may feel driven to ask questions or seek new knowledge to meet normative performance standards or to outperform others. With the addition of this measure, one could also examine the ability of the CIAC to differentiate between children who seek to attain knowledge for their own benefit (i.e., mastery orientation motivation) and those who seek to perform well (i.e., performance approach motivation). Due to practical limitations of the current study, however, we did not include such measures. Future studies may investigate these possible relations further.

Future research could also examine which components of the CIAC are most important for fostering epistemic curiosity in children. For example, it could be that children's curiosity is best fostered by first attending to their perceived value of epistemic curiosity for their own learning (*Personal Inclination*) before attending to their perceived social norms or behavioral control (for research about such considerations, please see Vogel & Wänke, 2016). In addition, further research into possible determinants of children's epistemic curiosity might also include children's

beliefs about the malleability of their own learning abilities through active learning engagement (e.g., Blackwell, Trzesniewski, & Dweck, 2007).

Importantly, future research should investigate the ability of the CIAC to predict children's epistemic behavior in school. Although we provided evidence for the predictive validity of the CIAC on the basis of motivational measures (in accordance with the Theory of Planned Behavior), the absence of behavioral data in the present study can be seen as a limitation. However, as others have argued as well (e.g., Engel, 2006), children may not exclusively 'express' their epistemic curiosity through *overt* behavior that is observable in the classroom, such as by verbalized question-asking or explanation-seeking behavior, but as much so through *covert* behavior, such as by engaging in curiosity-driven thought while working on study assignments. Therefore, a multi-method approach to assessing children's (developing) epistemic curiosity behavior is needed that includes specialized classroom observation and in-depth interviews with children and teachers to measure the range of possible behaviors and thought-processes associated with children's epistemic curiosity in formal education settings. To our knowledge, however, no such validated measurement instrument yet exist. For this reason, we were unable to reliably assess children's epistemic curiosity behavior as part of the study. Future directions of curiosity research should thus include the development of such behavioral measures.

Conclusion

With this study, we hope to contribute to research on children's epistemic curiosity in formal education settings by broadening the scope of research beyond children's curiosity *behavior* to children's *images* of and *attitudes* towards being curious learners. In our view, research on children's attitudes towards curiosity will lead us to a more complete understanding of why children behave at school the way they do. Our findings provide empirical evidence that supports propositions in previous research (e.g., Claxton & Carr, 2004; Engel, 2006; Engel & Randall, 2009) that children's epistemic curiosity in school may only be moderate at best and seems to decline throughout primary school.

In line with this research, we believe that teachers should explicitly cultivate a positive classroom climate in which children are inspired to adopt epistemic images of curiosity and are made aware of the value of asking curious questions for their own learning and for knowledge acquisition or innovation in general. For example, teachers may lead group discussions among children to expose their pre-existing

narrow, naïve and negative perceptions about the epistemic value of curiosity (e.g., Abd-El-Khalick, 2012; Deng et al., 2011). In addition, teachers should pay attention to children's potential negative opinions or their fears of their classmates' judgments and strive to cultivate a sense of pleasure and pride in asking epistemic questions or posing new ideas (Post & Walma van der Molen, in press). Simple reward systems could further convey to children that their epistemic curiosity is part of the assessment of their learning in school. The CIAC questionnaire may provide researchers and educators with a useful measurement tool to evaluate the effectiveness of such pedagogical interventions.

Chapter 4

4 Effects of a longitudinal school development program on primary teachers' attitudes towards inquiry teaching and their inquiry teaching practices

Abstract

The present paper describes a two-year school improvement program in which the complete school staffs of six Dutch primary schools were trained to integrate inquiry-based pedagogy into daily school practice. A delayed treatment pretest-posttest control group design was used to investigate the effects of the program on teachers' attitudes towards inquiry teaching and their inquiry teaching behavior. In addition, differences in these program effects were explored based on differences in school leadership between the schools. Therefore, program effects were examined at the treatment level and at the individual school level by means of attitude questionnaires and school principal, teacher, and pupil interviews. Results indicate that the program significantly improved teachers' attitudes and inquiry teaching behavior of the first treatment group and that teachers' improved practices persisted to one year after training. However, school leadership remained largely moderate throughout the program and appeared to explain little variation in teachers' inquiry teaching development among the participating schools. In addition, results reveal positive but limited effects of the same program on the attitude and behavior development of the delayed treatment group. Findings provide insight into the factors that may foster or hinder primary teachers' adoption of inquiry-based pedagogy. Recommendations are given for the further development of inquiry-focused school development interventions.

Introduction

There is growing international emphasis on the implementation of inquiry-based pedagogy in primary education (Lucas, Claxton, & Spencer, 2013; OECD, 2015; Pellegrino & Hilton, 2012). Inquiry-based pedagogy expands teachers' traditional use of teacher-directed instruction with student-centered approaches to teaching that stimulate pupils' own investigation of school subject matter (Claxton, 2007; Stuckey, Hofstein, Mamlok-Naaman, & Eilks, 2013). However, helping whole primary school teams[1] to acquire the knowledge, skills, and attitudes to meet these new education standards is a complex enterprise, which requires extensive teacher enhancement programs (Capps, Crawford, & Constas, 2012; Sawyer, 2011). Unfortunately, such programs are still relatively scarce (Syer, Chichekian, Shore, & Aulls, 2012).

Most research on the above topic has been limited to: (1) short-term professional development aimed at developing primary teachers' knowledge, skills or attitudes for mastering only few inquiry teaching tasks or for the implementation of prescribed inquiry-based (science) lessons (Slavin, Lake, Hanley, & Thurston, 2014), which leaves teachers largely unprepared to encourage pupils' inquiry as an integral component of their daily teaching practice (see Claxton, 2007); (2) professional development of subsets of motivated teachers, without providing insight into the impact of whole-team professionalization on school-wide practice reform (Thurlings, Evers, & Vermeulen, 2015); (3) program evaluations based on few and general teacher variables, often examined only at the treatment level (i.e., across treatment schools), without providing insight into possible differences in program effects between individual schools by varying school leadership and organization (for general theory about this issue, please see Berliner, 2002); (4) program evaluations immediately after training completion, which provide little insight into the degree to which schools managed to sustain their improved practices after training (for general recommendations on this issue, please see Timperley, 2008); and lastly (5) non-experimental research (e.g., expert opinions, case studies, cross-sectional studies),

---

[1] In the remainder of this chapter, we refer to 'school teams' as comprising *all* school staff (all teachers and the school principal). Please note that, in The Netherlands, primary teachers teach all school subject matter themselves (including science). Schools may decide how the subjects are taught and what teaching materials to use. The Dutch primary school curriculum emphasizes numeracy and literacy education, but also includes social and environmental studies (e.g., science, geography, history, biology, citizenship).

which offers limited empirical evidence for the efficacy of the professionalization treatments under investigation. Given the above, little is currently known about what it takes to help primary schools integrate inquiry-based pedagogy into their school curriculum and which teacher and school factors might foster or hinder teachers' professional development in this regard.

The present study

To help fill the above void in the literature, the present paper describes the effects of a two-year school improvement program in which six Dutch primary school teams were trained to adopt inquiry-based pedagogy. These effects were examined based on a wide array of teacher and school leadership variables at both the treatment level and the individual school level. We examined these effects longitudinally (one pretest and two posttests), and compared the results of our treatment to school teams that did not receive the same training. Clearly, such longitudinal school development research is complex, expensive and labor-intensive (Desimone, 2009). However, such an approach may advance our understanding of what types of support primary school teams need to adopt inquiry-based pedagogy.

Below, we first provide a working definition of inquiry-based pedagogy for this study, since many different definitions exist in the literature. Next, we propose what teaching and school leadership qualities are regarded important for inquiry-focused teaching reform in the literature. Lastly, we describe our intervention study design, hypotheses and research questions.

Defining inquiry-focused teaching for this study

Scientific descriptions of inquiry-based learning can be distinguished into the following three goals (e.g., Dewey, 1910; National Research Council, 2000, 2012; Osborne, 2014): (1) understanding how inquiry by scientists proceeds (i.e., learning *about* inquiry); (2) being able to perform inquiry (i.e., learning *to* inquire); and (3) constructing an understanding of (science) subject matter by inquiry (i.e., learning *by* inquiry). In the present study, we focus on the latter goal, more specifically, on inquiry pedagogy aimed at helping pupils *use* inquiry as a general strategy to study school subject matter (see Claxton, 2007). Research suggests that constructivist approaches to learning, such as inquiry-based learning, may help pupils develop a meaningful and integrated understanding of school subject matter, as they are actively involved in the

construction of their own ideas, solutions and explanations (Bruner, 1961; Papert, 1980; Syer et al., 2012; Walan, McEwen, & Gericke, 2016).

Inquiry encourages pupils' own inquisitive ideas and questions to emerge from their own study of school topics and research projects and, in turn, challenges teachers to be responsive to pupils' ideas and questions by elaborating or extending these through group discussions, design experiments, or new investigations (Bennett, Lubben, & Hogarth, 2007). This way, inquiry stimulates pupils to curiously, creatively, and confidently study (novel) school subject matter mostly on their own (Claxton, 2007), to make productive connections between subject domains (Heywood, Parker, & Jolley, 2012), and to develop a more meaningful and integrated understanding of the subject matter generally taught in school (Walan et al., 2016).

Inquiry-based teaching is thus not limited to science content alone, but allows pupils' inquiry to take place across different subject domains, including traditional domains such as English, geography, and history (see NGSS, 2013; OECD, 2015; Stuckey et al. 2013). It expands the role of the teacher to go beyond that of following lesson methods chapter by chapter, to the role of creatively *enhancing* lesson activities with opportunities for pupils to inquire, such as solving real-life problems, conducting experiments, designing and evaluating solutions (Osborne, 2014).

Teacher and school principal qualities for inquiry-focused teaching practice

Over the past two decades, research has recommended several changes to introduce inquiry-based pedagogy, especially in the domains of scientific literacy (e.g., Lumpe et al., 2000), science education (e.g., Osborne, 2014), creativity (e.g., Sawyer, 2015), and twenty-first century learning (e.g., Pellegrino & Hilton, 2012). The present study is based on the contention that, to achieve sustained inquiry-focused teaching practice in primary education, it is crucial for primary teachers to develop positive attitudes towards inquiry-based teaching. Studies indicate that teachers' inquiry teaching behavior is profoundly shaped by their attitudes towards inquiry-based teaching (Osborne, Simon, & Collins, 2003; Lumpe, Czerniak, Haney, Beltyukova, 2012). However, most primary teachers find it difficult to teach school subject matter through inquiry, because they themselves lack sufficient familiarity with inquiry (Ricketts, 2014). As a consequence, most primary teachers hold negative attitudes towards inquiry teaching and thus often shy away from inquiry-based lesson activities (Jarvis & Pell, 2004). Encouraging inquiry teaching practice in primary schools thus

calls for teachers to become sufficiently familiar with enhancing opportunities for pupils to conduct inquiry and to improve their own attitudes toward inquiry teaching.

In addition, the school improvement literature promotes several factors at the school leadership level that are important for successful practice reform (e.g., Thoonen, Sleegers, Oort, Peetsma, & Geijsel, 2011). The impact of school leadership on organizational performance and teachers' professional development in general has so far proved to be only modest (Muijs, 2011). Nevertheless, we presumed school leadership aimed at fostering teachers' implementation of inquiry-based teaching practice to be important. We thus set out to explore this possible relationship in the present study as well.

Below, we used common themes in the above literatures to list possibly relevant qualities of inquiry-oriented teachers and school principals.

Teachers' attitudes towards inquiry teaching. Attitude can be defined as the evaluative beliefs a person has about a particular behavior in a certain context in terms of favorable or unfavorable features (Ajzen, 2001; Ajzen & Fishbein, 1980; Eagly & Chaiken, 1993), such as the benefit of the behavior to the person or the pleasure of engaging in the behavior. These beliefs determine the intention of the person to enact the behavior, either covertly or overtly, when sufficient opportunity arises to do so (Ajzen, 2001).

A recent effect study by Van Aalderen-Smeets and Walma van der Molen (2015) revealed that teachers' self-efficacy and context dependency beliefs about inquiry-based (science) teaching are essential components of teachers' professional attitudes, as both components showed to significantly affect teachers' frequency of inquiry-focused teaching. According to the Theory of Planned Behavior (Ajzen & Fishbein, 1980), teachers' self-efficacy and context dependency beliefs can be regarded as two separate, but closely related, components of teachers' perception of behavioral control (for a detailed description of the theoretical underpinnings of these concepts, please see Van Aalderen-Smeets & Walma van der Molen, 2011).

Teachers' self-efficacy can be understood as teachers' perceived competency (Bandura, 1997). Teachers who hold positive self-efficacy beliefs about inquiry teaching are more likely to adopt inquiry into their practices, put more effort into planning and delivering their inquiry-oriented lessons, and better persevere in the face of difficulty than teachers with less positive self-efficacy beliefs (Osborne et al., 2003). Studies reveal that many primary teachers have naturally low self-efficacy about teaching (science) through inquiry (Tosun, 2000). The improvement of teachers' self-efficacy has therefore been recognized as a critical component of

inquiry-focused teacher professional development (Carleton, Fitch, & Krockover, 2008).

Teachers' context dependency is defined as teachers' perceived dependency on contextual factors for teaching school subject matter through inquiry, such as time and resources available (Ford, 1992). Studies reveal that many primary teachers hold misconceptions about the context factors that they believe impede teaching through inquiry (Lumpe et al., 2012). Many primary teachers who hold positive self-efficacy beliefs about inquiry teaching, still refrain from adopting such practice because they believe that it requires prescribed teaching materials or competes with national education standards. Thus, the improvement of teachers' perceived context *independency* is considered vital to make teachers feel inclined to perform as inquiry-oriented teachers (Van Aalderen-Smeets & Walma van der Molen, 2015).

As much as primary teachers need to hold positive attitudes towards inquiry teaching in general, we believe that teachers should also possess positive beliefs about *creatively enhancing* their usual teaching methods with inquiry teaching methodology (for theory about such beliefs, please see Thurlings et al., 2015). To encourage pupils' inquiry, teachers must feel willing to incorporate opportunities into their regular lesson activities for pupils to conduct inquiry. Therefore, for example, teachers can revise typical teacher-directed lesson activities into more open-ended lesson activities or to encourage pupils' inquiry during lesson activities by means of questioning techniques. However, studies show that most primary teachers rather prefer to follow lesson books chapter by chapter and, when deciding to incorporate inquiry-based lesson activities, seek comfort in using prescribed inquiry lessons and materials instead (Jones & Eick, 2007). As such, we contrast teachers' positive perceptions of *creative* lesson design with teachers' predominant perceptions of *prescribed* lesson design. We conceptualize both concepts as two separate, but related, components of teachers' perceptions of lesson design.

Teachers' inquiry teaching behavior. Inquiry teaching practice not only requires teachers' *willingness* to encourage pupils' inquiry, but as much so requires their *ability* to do so. Being able to create a stimulating epistemic classroom culture for inquiry is considered one of the most important, if not the most important, prerequisite for stimulating pupils' inquiry (Carr & Claxton, 2004; Claxton, 2007).

In particular, research increasingly promotes the cultivation of primary pupils' curiosity-driven question-asking and explanation-seeking behavior (Engel, 2015). According to Post and Walma van der Molen (2018a, 2018b), such a classroom culture can be characterized by pupils' positive beliefs about the value of question-asking and

explanation-seeking for their own learning, about the pleasure to engage in such behavior, and about the appreciation they receive from their teachers and peers for showing such behavior. Post and Walma van der Molen (2018b) propose that such positive beliefs can be understood in terms of pupils' attitudes towards epistemic curiosity. To cultivate pupils' epistemic curiosity, the authors thus propose that inquiry teaching should focus on fostering pupils' attitudes towards epistemic curiosity. Through conducting reflective group discussions, for example, teachers can make pupils aware of their (implicit) beliefs about epistemic curiosity and subsequently challenge them to adopt more positive beliefs about these matters, such as by demonstrating to pupils the use of epistemic curiosity as a strategy for learning or the pleasure of engaging in curiosity-driven inquiry (for examples, see Post & Walma van der Molen, 2018a).

At the same time, teachers should stimulate pupils' inquiry by encouraging their 'higher-order' thinking during their study of school subject matter (Osborne, 2014). Higher-order thinking involves the synthesis, evaluation or analysis of information in order to come to new solutions, ideas or questions (Anderson & Krathwohl, 2001; King, Goodson, & Rohani, 2011). Questioning techniques are considered one of the primary means by which teachers can stimulate pupils' higher-order thinking (Marzano, Pickering & Pollock, 2001). For example, teachers can rephrase traditional content-oriented study assignments into more ill-structured, open-ended assignments that allow for multiple solutions and interpretations (Lederman & Abell, 2014; Miri, David & Uri, 2007). In addition, teachers can ask pupils different types of questions during lesson activities, such as aimed at making productive connections between subject domains or to consider alternative answers or explanations to presented questions and problems (Claxton, 2007; King et al., 2011).

Because one of the aims of inquiry is to engage pupils in studying novel and complex subject matter, teachers should also aim at cultivating pupils' mindset to persist their inquiry in the face of learning difficulty (Claxton & Carr, 2004). Such learning "resilience" can be understood as pupils' implicit belief that their own inquiry ability is malleable and can be largely improved through deliberate and sustained effort (Blackwell, Trzesniewski, & Dweck, 2007; Dweck, 2000). According to Dweck (2000), and others, such a 'growth mindset' determines a wide range of achievement motivations related to inquiry learning, such as how pupils approach learning, respond to challenges, and cope with setbacks during learning (see also Burnette, O'Boyle, VanEpss, Pollack, & Finkel, 2013). Teachers can positively develop pupils' growth mindset. For example, studies show that teachers' praise of pupils' learning

effort rather than their learning ability, conveying to pupils the joy of tackling challenging learning tasks, and highlighting their progress and effort, positively impacts pupils' academic performance (e.g., Blackwell et al., 2007).

Lastly, inquiry teaching practice requires teachers' daily use of the above-described inquiry teaching methodology (see Claxton, 2007). Furthermore, every day, teachers should provide pupils substantial time to conduct inquiry, such as for coming up with their own inquisitive questions and ideas and to answer or test those (at a later time) during class (Biggers & Forbes, 2012). To stimulate pupils' inquiry further, teachers should present school subject matter to pupils in an integrated, cross-curricular manner, such as by overarching socio-scientific themes (e.g., energy, transport, climate). This maximizes the likelihood of pupil discussions and investigations into the interrelatedness of school subject domains (Heywood et al., 2012). In addition, teachers should regularly evaluate their inquiry teaching by the above quality standards.

School principals' leadership behavior. As discussed above, teachers play an integral role in determining the success of school practice reform. However, the complex task of integrating inquiry-based pedagogy into daily school practice falls upon the responsibility of the schools to work out themselves (Lumpe et al., 2000; Osborne & Dillon, 2008). School principals are assumed to play an important role in this regard, as they are generally assigned the task to create favorable conditions for teachers to realize and sustain their improved practices (Leithwood, Harris, & Hopkins, 2008; Murphy & Seashore, 2018).

Therefore, school principals should set and communicate clear policy on the direction and expectations of teachers' inquiry-focused practice reform. By actively involving themselves with teachers' reform efforts and by learning from teachers' experiences with adopting inquiry-based pedagogy, they can formulate performance goals at the school level and decide on strategies to achieve these goals (Murphy & Seashore, 2018). In parallel, school principals should regularly foster purposeful discussions among school team members about the evaluation of their school performance goals, such as during monthly staff meetings.

In addition, research indicates that school principals should actively encourage and facilitate teacher collaboration on practice reform (Thoonen et al., 2011). School teams with highly collaborative members better facilitate the adoption of new professional routines than schools without much teacher collaboration and, consequently, show improved individual and organizational performance (Moolenaar, Sleegers, & Daly, 2012). Therefore, school principals should encourage

teachers to collectively experiment with adapted classroom practice (e.g., sharing ideas, questions, and challenges) and provide them sufficient time and resources during regular school hours for such collegial collaboration (Hargreaves, Lieberman, Fullan, Hopkins, 2010).

Lastly, at the cultural level, the vision and norms promoted by a school principal are assumed to have a positive effect on practice reform as well (Leithwood, Harris, & Hopkings, 2008). School principals who communicate a clear vision about inquiry-based pedagogy in their school, who prioritize school norms on classroom experimentation and teacher collaboration, and emanate to their school staff a sense of collective efficacy for practice reform, will likely be more successful with implementing inquiry-based teaching practice than school principals who do not (Thurlings et al., 2015). School principals may thereby function as influential role models for the teachers.

The intervention

A nine-months teacher enhancement course was developed for six Dutch primary school teams aimed at developing teachers' knowledge, skills and attitudes for inquiry-based teaching practice, and implementing and sustaining such practice in the entire school organization. In the development of this course, we attempted to meet the aforementioned quality standards of inquiry-based teaching and school leadership as found in the literature.

Because improving teachers' attitudes was considered essential for practice reform, one of the main purposes of the program was to improve teachers' attitudes towards teaching inquiry by means of attitude-focused teacher professional development. Therefore, the intervention included the complete six-months attitude-focused teacher professional development course by Van Aalderen-Smeets and Walma van der Molen (2015), which aims at improving primary teachers' attitudes toward teaching (science) through inquiry. In a recent experimental study, this course was found to significantly improve teachers' professional attitudes towards inquiry-focused teaching practice, teachers' personal attitudes towards science, and teachers' frequency of inquiry-focused teaching (Van Aalderen-Smeets & Walma van der Molen, 2015). We extended this course with a newly developed, three-months didactical training course aimed at developing teachers' knowledge and skills for stimulating and assessing pupils' inquiry during usual classroom practice. School principals were required to attend all course meetings and were encouraged to act as

inquiry-oriented leaders throughout the program (but received no inquiry-focused leadership training).

Hypotheses and research questions

Teachers' attitude development. As previously described, the results of Van Aalderen-Smeets and Walma van der Molen (2015) indicated that their attitude-focused teacher training course affected positive changes in participants' self-efficacy and context dependency beliefs that relate to inquiry teaching. Because the current intervention included the same complete course, we expected that:

H1: Participants' self-efficacy belief mean scores would significantly increase, while their context dependency belief scores would significantly decrease as a result of participating in the program, compared to a control group of teachers who did not participate in the training.

In addition, our current school development program included explicit professionalization activities aimed at motivating participants to creatively enhance their regular teaching methods with inquiry-based pedagogy. Therefore, we expected that:

H2: Participants' creative lesson design belief mean scores would significantly increase, while their prescribed lesson design belief mean scores would significantly decrease as a result of participating in the program, compared to a control group of teachers who did not participate in the training.

Teachers' behavior development. Based on the findings by Van Aalderen-Smeets and Walma van der Molen (2015), we expected teachers' attitude improvement to coincide with increased inquiry teaching behavior. We thus hypothesized that:

H3: Participants' inquiry teaching frequency, inquiry teaching duration, and inquiry teaching evaluation would significantly increase as a result of participating in the program, compared to a control group of teachers who did not participate in the training.

In addition, the didactical training was aimed at improving teachers' inquiry teaching *quality* as well. In particular, teachers were trained to stimulate pupils' attitudes

towards epistemic curiosity, their higher-order thinking, and their growth mindset. The teachers were also trained to present school subject matter in a more integrated, cross-curricular, manner and to regularly evaluate their (developing) inquiry teaching quality. Given these aims, we expected that:

H4: Participants' scores related to the stimulation of pupils' attitudes towards epistemic curiosity, higher-order thinking, and growth mindset, as well as participants' cross-curricular inquiry teaching and their inquiry teaching quality evaluation, would significantly increase as a result of participating in the program, compared to a control group of teachers who did not participate in the training.

Lastly, attitude research indicates that changes in attitude are linked to subsequent changes in behavior and that this change is relatively stable (Vogel & Wänke, 2016). Thus, if teachers' attitudes towards inquiry teaching would in fact improve as result of the attitude training and, consequently, positively affect their inquiry teaching behavior, we expected teachers' improved inquiry teaching behavior to persist to the end of the program (see also Thurlings et al., 2015). In the present study, teachers' inquiry teaching behavior was mostly measured by qualitative methods. Thus, we asked the following research question:

RQ1: To what extent do participants sustain their inquiry teaching behavior from immediately after course completion to the end of the research program?

School principals' behavior development. Although the present intervention did not include leadership training for the participating school principals, we wondered whether the school principals would become more focused in their inquiry-focused leadership as a result of attending all course meetings. Based on our literature review, we distinguish several important features that characterize effective school principals' leadership behavior in this regard, namely communicating clear school policy (i.e., school improvement goal-setting, school improvement goal evaluation), facilitating teacher collaboration (i.e., inquiry teaching support, inquiry teaching adaptation support), and creating a positive school culture on adopting inquiry-based pedagogy (i.e., school improvement priority, school improvement motivation). With this study, we asked two research questions:

RQ2: To what degree do school principals improve their school improvement goal-setting, school improvement goal evaluation, inquiry teaching support, inquiry teaching adaptation support, school improvement priority, and school improvement motivation during the intervention?

RQ3: What of the school leadership factors might differentiate the schools that showed significant improvement in teachers' inquiry teaching development from those that benefitted significantly less?

Method

Design

A quasi-experimental pretest-posttest design with a delayed treatment group was used to assess the effects of the school development program (for a rationale of the use of such designs, please see Slavin, 2007). This 'staggered' intervention design allowed all the schools to receive training. To this end, the schools were randomly assigned to one of two treatment groups. The first treatment group received the intervention in the school year 2014–2015. The second group received the same intervention one year later, in the school year 2015–2016. All participants of all six schools were assessed prior to (T1, October 2014), during (T2, June 2015), and after the research project (T3, June 2016). Please see Figure 4.1 for a schematic overview of the research design.

In the first phase of the study (school year 2014–2015), we examined the effects of the intervention by means of a straightforward two-group pretest-posttest design (T1 vs. T2), where the first treatment group served as the intervention group and the delayed treatment group served as the control group (see Figure 4.1, solid box *A*). In the second research phase (school year 2014–2015), we examined the effects of the same (delayed) intervention for the delayed treatment group. This was done in two ways: (a) by means of a one-group double pretest-posttest design (T1 vs. T2 vs. T3) with only the delayed treatment group (see Figure 4.1, dashed box *B1*) and (b) by comparing the posttest results of the delayed treatment group at T3 to the posttest results of the first treatment group at T2 (see Figure 4.1, dashed ellipse box *B2*). At the end of phase two, we also investigated the degree to which the first treatment group had sustained their scores by means of a one-group longitudinal comparison of results over T1, T2, and T3 (see Figure 4.1, large dashed box *C*). We examined all of

the above intervention effects in two consecutive steps, namely first at the treatment level and second at the individual school level.

In addition, we invited three additional primary schools to serve as a comparison group of naive schools at T1. Research indicates that teachers' interest in inquiry teaching is related to their attitudes towards science and inquiry (Lumpe et al., 2012). However, because our treatment schools did not have extended experience with inquiry teaching yet, we did not expect that the teachers in these six schools would already hold positive attitudes towards teaching inquiry to begin with. To test this assumption, we added the naive control group at the beginning of our longitudinal study (T1) for a baseline score comparison of teachers' attitude scores (see Figure 4.1, dotted box *X*).

Schematic overview of the staggered research design

	Phase 1		
		Phase 2	
	October 2014	June 2015	June 2016
First treatment group	T1	T2	T3
Delayed treatment group	T1	T2	T3
Control group	T1	–	–
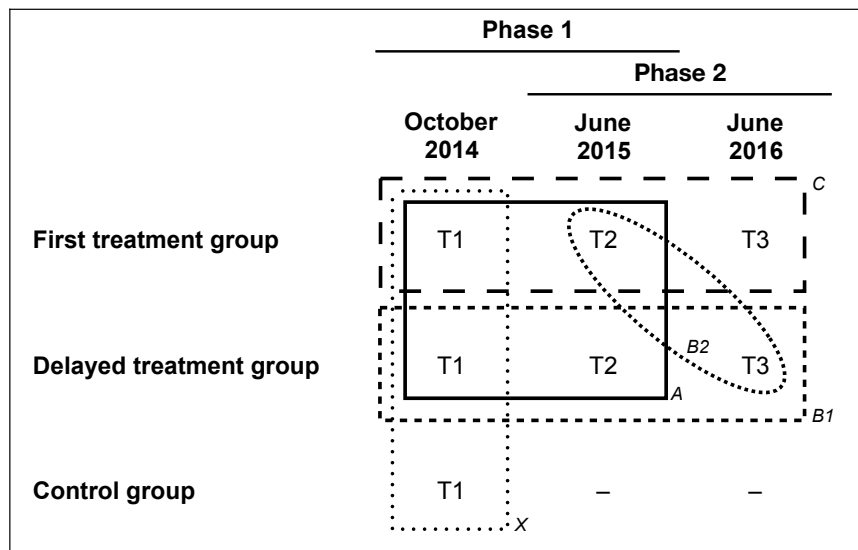
Figure 4.1. Schematic overview of the staggered research design; T1, T2, and T3 represent the consecutive measurement times. Box *A* (solid outline) represents the experimental pretest-posttest design with the first treatment group as the experimental group and the delayed treatment group as the control comparison group. Box *B1* (dashed outline) represents the one-group double pretest-posttest design of the delayed treatment group. Ellipse *B2* (dashed outline) represents the comparison of the posttest results of the delayed treatment group at T3 to the posttest results of the first treatment group at T2. Box *C* (large dashed outline) represents the one-group longitudinal results comparison of the first treatment group. Box *X* (dotted outline) represents a baseline score comparison of teachers' attitude scores before treatment at T1.

Sample

Six Dutch primary schools (K-8, i.e. age 4–12) enrolled in the study that were willing to commit themselves with their whole school team to our longitudinal intervention and measurements. We required schools to pay for enrolling into the first part of the training course, because we employed a trained teacher-educator to train both treatment groups. This procedure also ensured the commitment of the schools to the longitudinal development program. Each school declared they had not received inquiry-focused teacher professional development in the past.

The six schools comprised a total of 120 teachers and 6 school principals at T1 (83% female, mean age of 41.7 years, *SD* = 11.7). The first treatment group comprised School 1 (*n* = 17 teachers), School 2 (*n* = 15 teachers), and School 3 (*n* = 29 teachers). The delayed treatment group comprised School 4 (*n* = 18 teachers), School 5 (*n* = 32 teachers) and School 6 (*n* = 9 teachers). During the program, every school year, four pupils (2 boys and 2 girls) were randomly selected from each 4th, 5th and 6th grade classroom per school to take part in pupil interviews. The pupil interviews served us to measure changes in teachers' inquiry teaching behavior during the school development program. In total, we interviewed 48 pupils in the first research phase and 48 pupils in the second research phase. During the two school development phases, 13 teachers dropped out due to unforeseen personal and organizational circumstances, leaving a total of 113 school team members who finished the program from start to finish.

The naive control group of schools was recruited from the same area as the treatment schools. These three schools also declared to not have received any inquiry-focused teacher professional development in the past. The control group was debriefed about the purpose of the study after the schools were assessed at T1. The three control group schools at T1 comprised 31 teachers and 3 school principals (82% female, mean age of 42.1 years, *SD* = 11.7).

All participants agreed to fill out our questionnaires and, in the case of the treatment schools, to be interviewed regularly during the extent of the school development program. Informed consent was obtained from the parents of all the pupils who were interviewed in the study in accordance with the ethical guidelines of our university. All schools agreed to not participate in any other formal teacher training courses during the two years of the study.

Measures

Teachers' attitudes towards inquiry teaching. Attitude data were collected before treatment and once after treatment using self-report questionnaires.

Thus, at T3, we did not collect attitude data among the first treatment group. The original questionnaire consisted of 21 pre-written attitude statements that measured respondents' perceptions of *Self-Efficacy*, *Context Dependency*, *Creative Lesson Design,* and *Prescribed Lesson Design* regarding their own inquiry teaching. Respondents could indicate to what extent they agreed or disagreed with each attitude statement by use of a five-point Likert scale, ranging from 'totally disagree' (score 1) to 'totally agree' (score 5). The complete questionnaire is listed in the Appendix A.

Respondents' perceptions of *Self-Efficacy* and *Context Dependency* were assessed by means of two subscales that were taken from the DAS questionnaire (for detailed information about the psychometric quality of the DAS questionnaire, please Van Aalderen-Smeets & Walma van der Molen, 2013). The *Creative Lesson Design* (six items) and *Prescribed Lesson Design* (8 items) subscales were newly developed for the purposes of this study and added to the questionnaire.

We tested the psychometric quality of the above attitude subscales on the basis of the T1 data. Because the questionnaire was partially new, we conducted exploratory factor analysis (EFA) with principal axis factoring and direct oblimin rotation to verify our hypothesized four-factor structure. EFA revealed that our hypothesized four-factor structure fitted the data best, explaining 32.31% of the total variance. Three items of the *Prescribed Lesson Design* subscale and two items of the *Creative Lesson Design* subscale were removed due to poor loadings, leaving a total of 16 items. The factor loadings of the remaining items were all sufficient, as indicated by loadings greater than .45 and no significant cross-loadings (Worthington & Whittaker, 2006). In addition, all inter-factor correlations were lower than .75, which indicated that each factor possessed sufficient discriminative power. The internal consistency of each subscale proved to be sufficient as well, as indicated by Cronbach's alpha values greater than .70. Please see Table 4.1 for the number of items, Cronbach's alpha values, and the correlation between the subscales for the pretest (T1) data.

Table 4.1

The number of items, internal consistencies, and correlations of the subscales at T1

	Number of items	Cronbach's alpha	Self-Efficacy	Context Dependency	Creative Lesson Design	Prescribed Lesson Design
Self-Efficacy	4	.79	1.00	$-.27^{**}$	$.48^{**}$	$-.18^{*}$
Context Dependency	3	.70	$-.27^{**}$	1.00	$-.36^{**}$	$.46^{**}$
Creative Lesson Design	6	.78	$.48^{**}$	$-.36^{**}$	1.00	$-.39^{**}$
Prescribed Lesson Design	3	.72	$-.18^{*}$	$.46^{**}$	$-.39^{**}$	1.00

*Correlation is significant at $p < .05$
**Correlation is significant at $p < .01$

Teachers' inquiry teaching behavior. To determine teachers' inquiry teaching behavior, we obtained teacher and pupil perception data by means of interviews at T1, T2, and T3. The delayed treatment group was interviewed only at T2 and T3. Although classroom observations are often used to measure teaching behavior in primary education (Goe, Bell, & Little, 2008), this method is labor-intensive and time-consuming (e.g., Hill, Charalambous, & Kraft, 2012). Alternatively, recent studies indicate that pupil perceptions can provide reliable and valid information about teachers' teaching behavior for research purposes, as pupils experience their teacher on an ongoing basis and thereby may provide reliable comparative judgements about (changes in) teachers' teaching behavior (Van der Scheer, Bijlsma, & Glas, 2018). Thus, we collected perception data from all the individual 4th, 5th, 6th grade teachers and from groups of four of these teachers' pupils at each treatment school. We only collected interview data from the teachers and the pupils in grades 4-6, because the present study was carried out in the context of a larger impact study in which we specifically examined pupils' inquiry performance development in these grades.

A pre-structured format of standardized interview questions was used to consistently ask each group of respondents about their perceptions about teachers' current inquiry teaching behavior in the same sequential order. We developed the interview questions on the basis of the course materials provided to the participants as part of the training course and that best exemplified the range of studied behaviors associated with each aspect of inquiry teaching behavior. Each question required the teachers and school principals to complete a self-report scale of three behavioral statements by use of a five-point Likert response-option scale.

In addition to each self-report scale, as part of each interview question, respondents were asked to illustrate their overall score on the scale with a personal account of a related experience (e.g., providing a recent example of stimulation of pupils' higher-order thinking). Please see the Appendix B for the complete interview format. We computed weighted sum-scores for each self-report scale by averaging respondents' scores on the three items that defined the scale. The internal consistency of each scale proved to be sufficient, as indicated by sufficient standard deviations (hovering around .7) and Cronbach's alpha values (> .70).

School principals' leadership behavior. School principal leadership behavior is typically measured by teacher perceptions (e.g., Robinson, Lloyd, & Rowe, 2008) or by school principal perceptions about their own behavior (e.g., Urick & Bowers, 2013). In the present study, we decided to measure teachers' and school

principals' individual perceptions at T1, T2, and T3 about school principals' leadership behavior for implementing inquiry-based pedagogy in their schools.

We developed a pre-structured format of standardized interview questions that best exemplified the range of leadership behaviors derived from our review of the school leadership literature. Each question required the school principals and the teachers (grades 4-6) to complete a self-report scale of three behavioral statements by use of a five-point Likert response-option scale. After each interview question, we asked the respondents to illustrate their overall score on the scale with a personal account of a related experience. Please see the Appendix C for the complete interview format. We computed weighted sum-scores for each self-report scale and found that the internal consistency of each scale was sufficient.

Training course

The teacher professional development course consisted of two consecutive parts: (1) a six-months attitude-focused teacher professional development course that was previously developed and tested by Van Aalderen-Smeets and Walma van der Molen (2015) and (2) a newly developed, three-months practical course by the present authors. The complete nine-months training program was structured in such a way that participants' attitudes, knowledge, and skills were gradually built up during the meetings. The complete school teams participated in both parts of the training. The entire training course adhered to the quality standards proposed by Wilson (2013) for effective inquiry-oriented teacher professionalization (e.g., active learning, practically relevant). Please see Appendix D for an overview of the main content of each course meeting.

Attitude-focused training. The attitude-focused course consisted of six meetings of 3 hours each (18 hours in total) that were scheduled with one month between each meeting. Participants spent about 40 hours in total on preparations and assignments. The course meetings combined lectures, workshops and discussions that challenged and raised participants' awareness about several important aspects of their attitude towards science and inquiry, including teachers' self-efficacy and context dependency beliefs. These assignments stimulated teachers to apply elements of the course content to their own classroom practices and to positively develop their attitudes towards teaching (science) through inquiry. For detailed information about the contents of the course, please see Van Aalderen-Smeets and Walma van der Molen (2015).

Didactical training. The second part of the course followed the same type of organization as the first course. The didactical training course consisted of three meetings of 3 hours each (9 hours in total) that were scheduled with three weeks between each meeting. Participants spent about 30 hours in total on preparations and assignments. Each participant received a reference guide, which contained theoretical background information about the training course and the didactical strategies required for stimulating and assessing pupils' inquiry.

The first meeting outlined to participants the relevance of stimulating pupils' positive attitudes towards curiosity, their growth mindset, and their higher-order thinking skills as vital components of pupils' inquiry learning capability. Participants learned about the pedagogical decision-making involved with stimulating these components in pupils during usual lesson activities, which concerned setting inquiry-focused lesson objectives and assessment criteria. In addition, participants learned to assess these objectives. We provided the teachers with worksheets to help record pupils' (developing) attitudes towards curiosity, growth mindset and higher-order thinking and to evaluate their own inquiry teaching quality on the basis of these results. Lastly, the teachers were asked to teach their pupils about the relevance, nature, and features of inquiry-based learning on the basis of the concepts discussed in the first meeting.

During the second meeting, participants evaluated the take-home assignment. They were then assigned the task to improve one of their own existing (typical) lessons (about any school subject matter) by use of the didactical strategies presented in the first meeting. During the meeting, the teachers from grades 1 to 6 were grouped in teacher pairs of closely related grade levels. Each teacher pair was asked to use the concepts from the first meeting to identify weaknesses in their chosen lessons and to propose suitable improvements to stimulate pupils' inquiry. Teachers were also instructed to plan their assessment of pupils' inquiry by use of the worksheets. The trainer guided teachers' work and provided feedback to all groups. At the end of the meeting, teachers were asked to implement their improved lessons in their own classrooms.

During the third and last meeting, each teacher pair presented their preparations, their implementation, and their evaluation of their 'inquiry-infused' lesson intervention to the school team and the trainer. For each presentation, all the school team members were stimulated to discuss possible improvements. The trainer guided these discussions, provided feedback, and encouraged continued professional

development. Lastly, the trainer stimulated the school team to consider opportunities for the further implementation of inquiry teaching into daily school practice.

Procedure

The attitude questionnaire was collectively administered by the first author to all school team members at their own school after regular school hours (the time duration was 30 minutes). The first author conducted the school principal, teacher, and pupil interviews at the school location during regular school hours. Around 10% of the total number of teachers in each treatment group worked in teacher pairs (working part-time). In these cases, the teacher who taught the class most days of the week was interviewed. For the pupil interviews, two boys and two girls were randomly selected from each 4th, 5th, or 6th grade classroom to be interviewed as a single unit. For each measurement time (T1, T2, T3), pupils were selected for the interview who had not taken part in the interviews before. The time duration of the school principal, teacher and pupil interviews was about 45 minutes. The interviewer consistently defined 'inquiry-based' teaching and learning to respondents at the start of each interview, as a few interview items used this textual phrase. The interviewer provided minimal encouragement to respondents' answers. Respondents' answers were processed by the first author on a computer during the interview and repeated to the respondent for clarification and confirmation.

The complete training course was organized at each school location. All school team members took part in the course meetings and completed the take-home assignments alongside their usual school activities. The attitude-focused training course was led by a teacher trainer with expertise in conducting and teaching inquiry and several years of experience with teaching this course to many primary school teams. This teacher trainer was not part of the research team. The consecutive practical course was developed by the present authors and led by the first author.

We note that one of the delayed treatment schools (School 5) felt discontent about the personal style of the first trainer after having completed the fourth meeting of the attitude-focused training course. To remedy the issue, the first author substituted the trainer at the fifth meeting and conducted the remaining sessions of the attitude-focused course in a somewhat adapted form.

Data analysis

Investigating participants' attitude development over time. The effects of the intervention on teachers' attitude development were analyzed by means

of one-way, GLM repeated measures MANOVA's using SPSS version 24. Therefore, teachers' self-efficacy, context dependency, creative lesson design, and prescribed lesson design scores were used as the dependent, within-group variables, while condition was used as the between-group variable. To gain insight into the effects of the intervention on each separate subscale, we performed post-hoc univariate analyses within each experimental group separately (intervention and control group).

Investigating participants' behavior development over time. To develop an integrated understanding of teachers' inquiry teaching behavior (development) during the intervention, we cross-examined teachers' perception scores on their own inquiry teaching behavior over time with their pupils' perceptions (for examples of such comparative analyses, please see Anderson, Leithwood, & Strauss, 2010; Hill et al., 2012). Similarly, we cross-examined school principals' and teachers' perception data concerning school principals' leadership behavior (development) during the intervention. Please note that the collection of the perception data among the pupils, teachers, and school principals was part of the treatment, since we expected that the interviews would likely foster respondents' reflection about their own behavior and thereby influence their attitude and behavior development. Changes in school principals', teachers' and pupils' perception scores were thus examined by one-group pretest posttest analyses at the treatment and the individual school level.

Preliminary data checks

Preliminary data checks showed that the distributions of teachers' attitude pretest scores (T1) on each attitude component met the statistical assumptions for analyses of variance. The percentage of missing attitude data among the participants in the first treatment group (2.11%), delayed treatment group (1.26%), and the control group (.55%) was acceptable.

Data checks also revealed that teachers' *Self-Efficacy*, *Context Dependency*, *Creative Lesson Design,* and *Prescribed Lesson Design* mean scores of both treatment groups at T1 were statistically equal, as indicated by independent *t*-tests per attitude component. In addition, we did not find significant differences on most of the attitude components between the treatment groups and the control group (Figure 4.1, box *X*). The *Prescribed Lesson Design* mean score of the control group was significantly greater than of the delayed treatment group (MD = .36), $t(82)$ = −2.28, p = .03). Overall, these results suggest that the initial attitude scores of the two

treatment groups were largely representative to the attitude score levels of the general teacher population in The Netherlands.

Results of research phase 1

Analyses at the treatment level

As previously described, we used a pretest-posttest control group design to test the effects of the intervention on the first treatment group (see Figure 4.1, box *A*). We describe the program effects first at the treatment level and second at the individual school level.

      Teachers' attitude development. Please see Table 4.2 for the mean pretest and posttest scores of the first treatment group, the delayed treatment group, and the control group on each attitude component per measurement time (T1, T2, and T3).

      To investigate the general effects of the intervention on participants' attitudes, a 2 (first treatment group vs. delayed treatment group) x 2 (T1 vs. T2) x 4 (*Context Dependency* vs. *Self-Efficacy* vs. *Prescribed Lesson Design* vs. *Creative Lesson Design*) repeated measures MANOVA was conducted with condition as the between-subjects factor, time as a within-subjects factor, and the four attitude components as dependent variables. MANOVA was performed on the weighted sum scores for all four attitude subscales, which were calculated by averaging respondents' scores on each set of items that defined the subscale. The test of within-subjects effects using Wilks' statistic revealed a significant overall interaction effect of time and condition, $\Lambda = .93$, $F(1, 89) = 6.99$, $p = .01$, $\eta^2 = .07$. This result shows that, across the four attitude components, the overall change in attitude was larger for the first treatment group compared to the delayed treatment group.

      Post-hoc univariate analyses revealed a significant but small interaction effect of time and treatment for teachers' *Self-Efficacy* scores, $F(1, 89) = 5.48$, $p = .02$, $\eta^2 = .06$. A paired *t*-test indicated a statistically significant increase in the *Self-Efficacy* scores of the first treatment group ($MD = .39$), $t(48) = -3.87$, $p = .00$, while the *Self-Efficacy* scores of the delayed control group remained statistically unchanged over time ($MD = .08$), $t(41) = .99$, $p = .33$.

      There was a statistically significant but small interaction effect of time and treatment for the *Context Dependency* component, where the *Context Dependency* scores of the first treatment group were significantly decreased in comparison to the scores of the delayed treatment group, $F(1, 87) = 7.00$, $p = .01$, $\eta^2 = .07$. Paired *t*-tests indicated a decline in the *Context Dependency* scores of the first treatment group (*MD*

= − .49), $t(48) = 5.08$, $p < .01$, while this effect was absent for the delayed treatment group ($MD = -$ .13), $t(41) = -1.34$, $p = ns$. Table 4.2 shows that the *Context Dependency* mean posttest scores of the first treatment group at T2 were moderate ($M = 2.54$, $SD = .91$).

We found a statistically significant and moderate univariate interaction effect of time and treatment for the *Creative Lesson Design* component, where the teachers in the first treatment group showed more increased *Creative Lesson Design* scores ($MD = .28$) than the teachers in the delayed treatment group ($MD = -$ .03), $F(1, 89) = 24.31$, $p = .00$, $\eta^2 = .12$. Paired *t*-test indicated a statistically significant increase in the *Creative Lesson Design* scores of the first treatment group ($MD = .28$), $t(48) = 4.93$, $p < .00$, while no attitude change was observed for the delayed treatment group ($MD = -$ .03), $t(41) = -.40$, $p = ns$.

No statistically significant univariate interaction effect of time and treatment was found for the *Prescribed Lesson Design* component, $F(1, 89) = 1.29$, $p = .26$, $\eta^2 = .01$. Paired *t*-tests indicated no significant changes in the scores of both groups for this component. Notably, the mean pretest score of both treatment groups for this component were already quite positive to begin with (T1) and had remained positive after training ($M = 1.98$; $SD = .75$).

In sum, the above results indicate significant small to modest effects of the training course on the *Self-Efficacy*, *Context Dependency*, and *Creative Lesson Design* attitude components of the first treatment group. Table 4.2 shows that the group achieved moderate to good attitude scores (T2). In preparation of our upcoming examination of teachers' behavior development in the upper-grade levels, and the presumed relation between their behavior and attitude development, we also note that the mean attitude scores of the upper-grade teachers were largely similar to the mean attitude scores of their teams (mean differences smaller than .4 across all attitude components).

Table 4.2
Means scores on each attitude component per experimental group

| | First treatment group | | | | | | Delayed treatment group | | | | | | Control group | |
| | T1 | | T2 | | T3 | | T1 | | T2 | | T3 | | T1 | |
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Self-Efficacy	3.23	.66	3.62	.71	–	–	3.26	.63	3.34	.64	3.47	.57	3.47	.84
Context Dependency	3.03	.79	2.54	.91	–	–	2.92	.91	2.79	.93	2.63	.95	3.17	.91
Creative Lesson Design	3.71	.58	3.99	.59	–	–	3.90	.65	3.87	.58	3.90	.65	3.71	.56
Prescribed Lesson Design	2.18	.63	1.98	.75	–	–	2.06	.65	2.00	.76	1.88	.73	2.40	.70

Note: Mean scores could range between 1 – 5; T1 = October 2014; T2 = June 2015; T3 = June 2016.

Teachers' behavior development. Next, we assessed the effects of the training course on the inquiry teaching behavior of the first treatment group by a one-group pretest posttest analysis from T1 to T2. Please see Figure 4.2 for the mean scores of the teachers and their pupils at the individual school level for each perception component per measurement time. Because of space limitations, we did not also plot respondents' scores at the treatment level.

Use of inquiry teaching. Wilcoxon Signed-Ranks Test revealed that the first treatment group had significantly and greatly improved its *Inquiry Teaching Frequency* (Z = −3.32; *p* < .01), *Inquiry Teaching Duration* (Z = −3.32; *p* < .01) from T1 to T2. At T2, the teachers indicated to teach inquiry on a 'daily' basis (*M* = 5.00; *SD* = .00), to spend 'half a day' on teaching inquiry at these times (*M* = 3.00; *SD* = .00), and to use inquiry across four school subject domains (*M* = 3.73; *SD* = .47) when teaching inquiry (i.e., excluding mathematics and literacy). As shown in Figure 4.2, their pupils perceived similar improvement for all three components (all *p*-values < .01).

Teachers' behavioral change was substantiated by various personal accounts of the teachers and their pupils during the interviews (T2). In particular, many teachers explained that the attitude training had inspired in them a renewed 'sense' of responsibility, urgency and flexibility to teach usual school subjects through inquiry. Their pupils described many lesson examples that illustrated teachers' (increased) use of inquiry teaching (e.g., assigning them the task to design a new Olympic stadium that could host both the winter and summer games). Several groups of pupils claimed that, in their eyes, their teachers seemed 'more interested in [their] ideas and questions during class'.

Inquiry teaching quality. Figure 4.2 shows that, prior to the first intervention (T1), teachers' *Stimulation of Pupils' Growth Mindset* mean score was quite positive and thus showed risk of a ceiling effect (we attempt to explain this high score at the individual school level in a later section). In addition, we note that pupils' scores on all teaching quality components were consistently lower than teachers' scores, as indicated by Mann-Whitney U test results (all *p*-values < .05).

After training (T2), teachers' *Cross-Curricular Inquiry Teaching Focus* (Z = −3.07; *p* < .01), *Stimulation of Pupils' Attitudes towards Curiosity* (Z = −2.99; *p* < .01) and *Stimulation of Pupils' Growth Mindset* (Z = −2.44; *p* < .05) showed significant and large improvement in the eyes of the teachers and had reached excellent score levels (mean scores above the 4-point level). Most teachers referred to the educational value of the training exercises as part of the didactical course and the reference guide

to enhance their lessons with inquiry. All teachers excluded the use of inquiry for teaching mathematics and literacy, because most felt that direct instruction fitted these lesson objectives better. However, teachers *Stimulation of Pupils' Higher-Order Thinking* mean score did not change during this time ($Z = -1.16$; $p = ns$), which had remained at a moderate score level ($M = 3.33$; $SD = .47$). Figure 4.2 shows that pupils' score patterns were very much alike teachers' score patterns. However, pupils' scores remained consistently lower than teachers' perception scores for all three components. Teachers' *Evaluation of Inquiry Practice* scores improved as well ($Z = -2.83$; $p < .01$) and achieved a good mean score ($M = 4.12$; $SD = .45$).

Most teachers supported their scores by various lesson examples, such as 'teaching pupils about the use of being curiosity-driven learners', 'teaching pupils to formulate higher-order research questions to discover new subject matter during their project work', and 'encouraging pupils' perseverance to complete difficult subject matter by praise and simple rewards'. Their pupils described similar teacher behavior, such as that '[their] teacher less often provided immediate answers, but asked questions when [they] did not quite understand how to solve problems on their own' or 'helped [them] to be come up with more exciting research questions to work on'.

Conversely, most teachers indicated that stimulating pupils' higher-order thinking turned out to be one of the most difficult components of their inquiry teaching practices. Many teachers stressed that the training course had helped them *prepare* lesson activities and projects that stimulated pupils' higher-order thinking (e.g., revising lesson objectives and assignments), but that the training mostly fell short in, as one teacher stated, 'learning to flexibly *respond* to pupils' spontaneous creative questions and ideas in the midst of ongoing lesson activities'. Therefore, a few teachers suggested the added value of follow-up coaching by the trainers in their own classrooms to further expand their practical knowledge and experience in this regard.
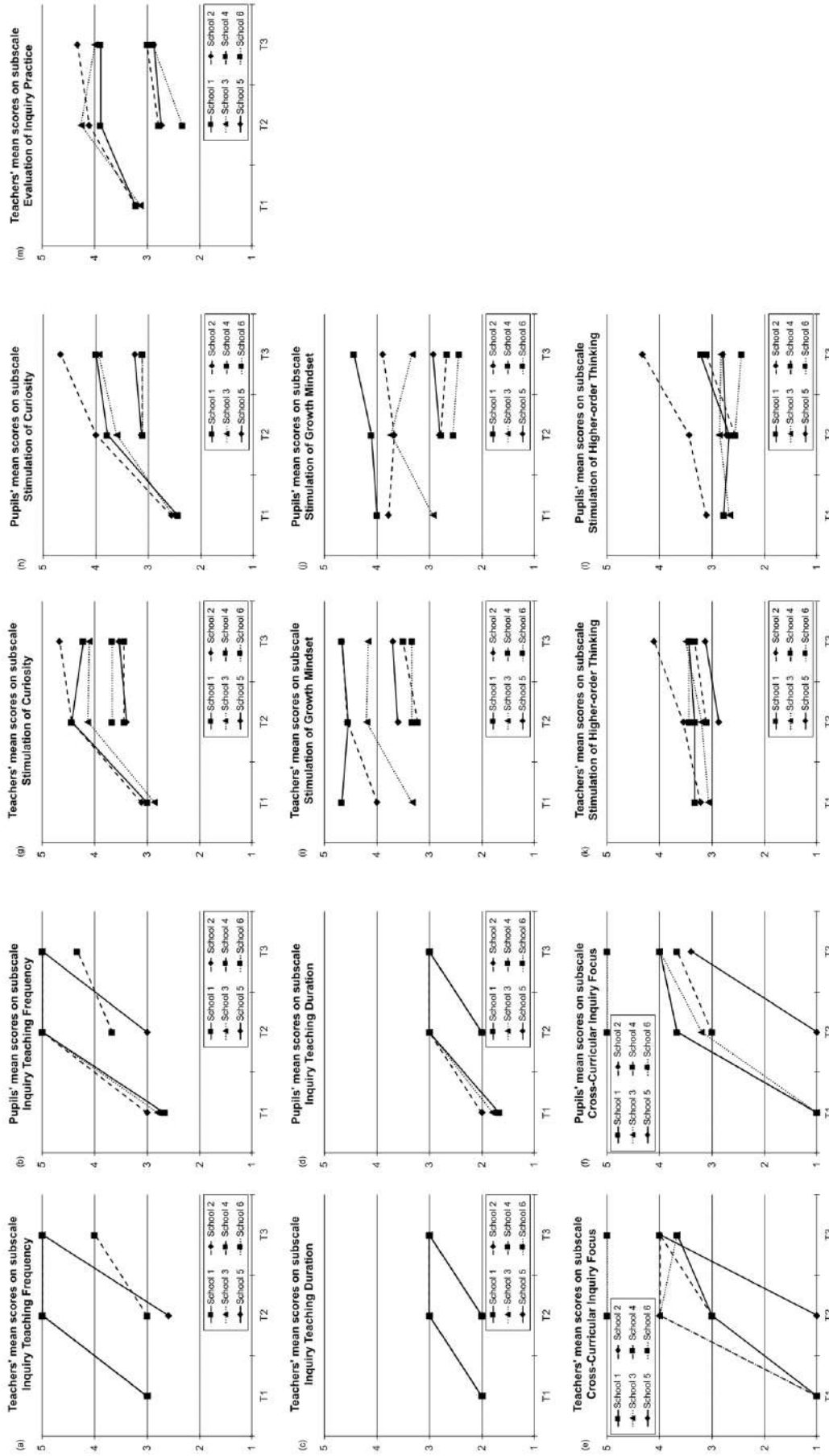
Figure 4.2. Teachers' and pupils' mean perception scores on inquiry teaching practice components at the school level per measurement time

School principals' behavior development. Lastly, we examined changes in school principals' leadership of the first treatment group during the first phase. Because of the small sample size of school principals in the first treatment group, we examined changes in school principals' scores by means of visual inspection only. Please see Figure 4.3 for the mean scores of school principals' and teachers' mean scores on each component of school organization per measurement time per treatment school.

We note that prior to the training (T1), school principals rated their *School Improvement Goals* (*M* = 4.11) ; *SD* = .38), *School Improvement Priority* (*M* = 5.00; *SD* = .00) and *School Improvement Motivation* (*M* = 4.33; *SD* = .58) by good to excellent scores, while their *School Improvement Evaluation* (*M* = 3.55; *SD* = .38), *Inquiry Practice Support* (*M* = 3.33; *SD* = .33), and *Inquiry Practice Adaptation* (*M* = 3.44; *SD* = .38) scores were moderate and thereby showed most room for improvement. Their teachers seemed to agree, as indicated by similar score patterns across all six school organization components. However, teachers appeared to rate their school principals' behavior consistently less positive than the school principals perceived their own behavior to be.

After training (T2), changes in school leadership were only perceived by the teachers in a few cases. In the eyes of the teachers, school principals' *School Improvement Goals* had significantly improved over time (Z = −3.00; *p* < .01), but still achieved only a moderate score level (*M* = 3.36; *SD* = .41). In addition, the teachers perceived a slight but significantly decline in *School Improvement Motivation* (Z = −2.12; *p* < .05).

School principals' and teachers' perception scores were supported by various personal remarks during the interviews. Most teachers expressed their appreciation for their principals' efforts to ensure a positive school climate for practice reform, but stated that they felt 'quite alone' in their teams to solve implementation problems and thus expected more support and guidance from their school principals. The school principals were generally more positive about these matters, indicating to have updated their school policy documents on inquiry teaching practice and discussing 'pertinent implementation issues' during monthly staff meetings. However, some school principals did indicate to experience difficulty leading teachers' implementation process. For example, two school principals stressed their difficulties with striking the right balance between 'instructing' versus 'assisting' teachers' adoption of inquiry teaching. One school principal stated that 'the success of [their] school improvement was ultimately dependent on teachers' intention to reform their

practice, not [his]'. In our view, this result suggests confusion among some teams about the responsible 'agents' of their practice reform. Both school principals indicated to desire specialized leadership training to learn to better foster teachers' implementation process.

Finally, we note that the above score patterns in school leadership of the first treatment group were not reflected by similar patterns in teachers' attitude scores, use of inquiry teaching scores, nor inquiry teaching quality scores between the schools. Despite of poor to moderate leadership, the teachers had shown significant improvement in their attitudes and inquiry teaching practices. This result thus suggests that school leadership on the implementation of inquiry teaching practice made little difference in teachers' attitude and behavior development.

Analyses at the school level

Thus far, we examined the effects of the first intervention at the treatment level. As described before in this paper, we were also interested in examining possible differences in program effects between the individual first treatment schools during phase one.

Teachers' attitude development. Please see Figure 4.4 for the mean scores of the individual first treatment schools (schools 1 to 3) for each attitude component at T1 and T2. Wilcoxon Signed-Rank test results revealed that participants' attitudes of School 1 and School 3 had slightly but significantly improved from T1 to T2, as indicated by decreased *Context Dependency,* increased *Self-Efficacy* and increased *Creative Lesson Design* mean scores (all p-values < .05). For School 2, we observed no significant changes in participants' mean attitude scores. However, after the training (T2), the mean attitude posttest scores of the first three treatment schools did not differ significantly, as indicated by non-significant Mann-Whitney U test results.
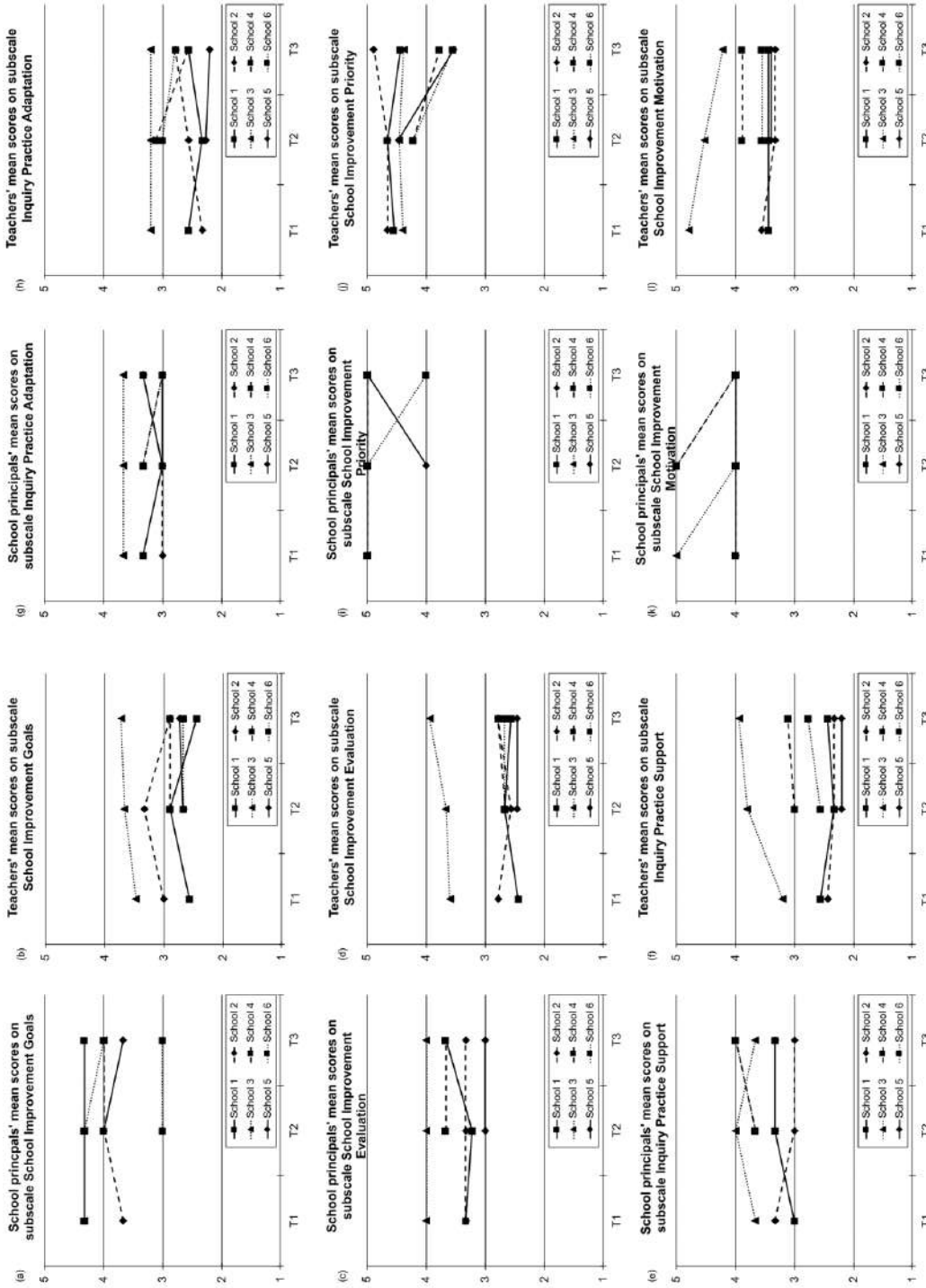
Figure 4.3. School leaders' and teachers' mean perception scores on school leadership components at the school level per measurement time
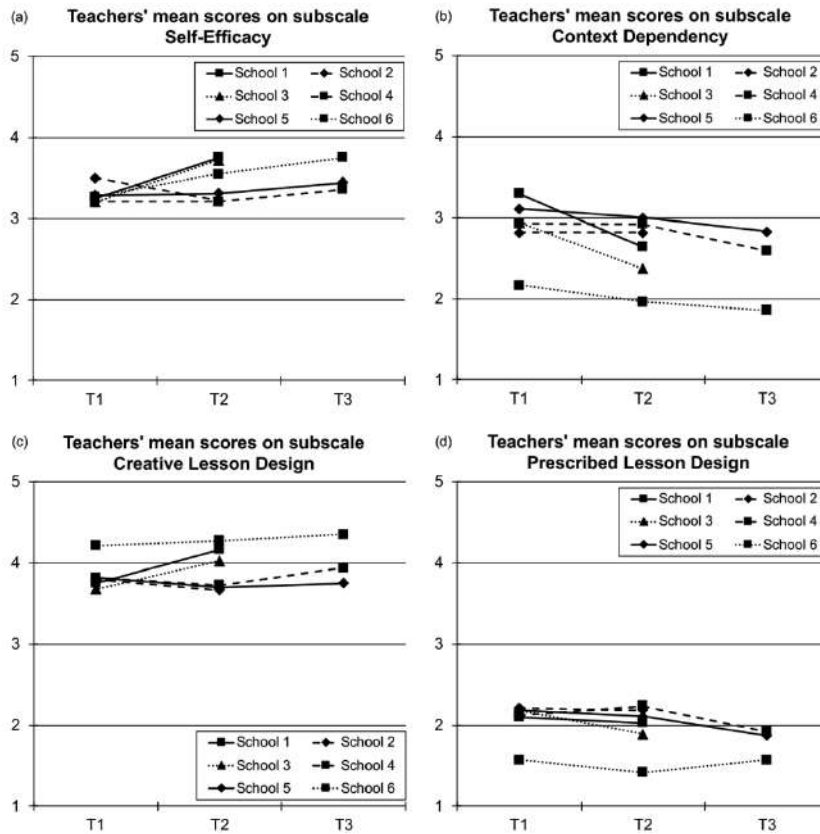
Figure 4.4. Teachers' mean scores on the attitude components at the school level per measurement time

In contrast to Schools 1 and 3, School 2 reported to both trainers to 'simply complete the entire training course' first and to work on the practical implications, challenges and opportunities for reforming their daily school practice afterwards to help the team focus. We believe that this 'serial' approach towards reform may have prevented the team from actively discussing their perceptions about school reform during the training, which may have in turn prevented participants' potential attitude development. This tentative explanation seems to be in part supported by the poor to moderate *Inquiry Practice Support, Inquiry Practice Adaptation,* and *School Improvement Motivation* mean scores of the teachers of School 2 during phase one, but do not seem to be entirely unique to School 2 (i.e., School 1's scores on these components were moderate too). Nevertheless, this finding underscores the importance of examining program effects at the individual school level.

Teachers' behavior development

Use of inquiry teaching. Figure 4.2 shows that teachers' *Inquiry Teaching Frequency* and *Inquiry Teaching Duration* mean scores were very much alike across the first treatment schools before (T1) and after (T2) the first intervention. Pupils' perceptions about these matters appeared very similar to teachers' perceptions at these times. This result indicates that schools' uniform attitude posttest scores seem to coincide with schools' uniform use of inquiry teaching after training, which thereby supports the positive relationship between teachers' attitudes towards teaching inquiry and their inquiry teaching behavior also at the individual school level.

Inquiry teaching quality. Figure 4.2 revealed that, in contrast to School 2 and 3, teachers' *Stimulation of Pupils' Growth Mindset* at School 1 was perceived to be exceptionally well-established by the teachers prior to the training (T1) and showed no improvement. This result was largely reflected by pupils' perception scores too. This reveals that the moderate increase of teachers' *Stimulation of Pupils' Growth Mindset* scores at the treatment level can be partly explained by School 1's high mean pretest scores (i.e., ceiling effect). The school principal and the teachers of School 1 explained their high pretest scores during the interview by earlier initiatives in their school that aimed at fostering pupils' growth mindset development.

In addition, Figure 4.2 reveals small but positive trends in teachers' *Stimulation of Pupils' Higher-Order Thinking* of School 2 on the basis teachers' perceptions. These positive trends are reflected by pupils' perception scores at this school as well. Conversely, teachers' perception scores at School 1 and 3 appear to have remained largely stagnant during the first phase, also in the case of their pupils' perceptions. This score pattern was not reflected in our qualitative interview data: at all three schools, the teachers expressed their difficulty with supporting pupils' spontaneous creative thinking during class. Altogether, these results thus reveal that, in contrast to our observations at the treatment level, that the intervention did seem to affect positive trends in teachers' *Stimulation of Pupils' Higher-Order Thinking* at School 2. Notably, this score pattern in inquiry teaching quality between the first treatment schools was not reflected by similar patterns in schools' mean attitude scores, as shown in our previous analysis. This result suggests that the development of teachers' inquiry teaching quality operated largely independent from teachers' attitude development.

School principals' behavior development. Across most school leadership components, School 3 appeared to show higher scores than School 1 and 2. Furthermore, we observe substantial improvement in school principals' *Inquiry*

Practice Support in the eyes of the teachers of School 3 during phase 1, while the teachers of School 1 and 2 perceived a moderate decline in scores on this component during this time. This result is supported by uniquely positive accounts provided by the teachers of School 3 about their school principals' leadership behavior compared to School 1 and 2. For example, several teachers of School 3 praised their school principal for cultivating a 'shared vision of inquiry teaching practice' among the schools staff (e.g., decorating areas in the school with slogans promoting pupils' inquiry, such as 'Beware, curious pupils at work!'), the facilitation of collegial support (e.g., teacher design team focused on the implementation of inquiry teaching), and frequent classroom visits to gather and provide feedback. The school principal of School 3 provided similar examples. In the case of *School Improvement Motivation*, however, School 3 did show a substantial decline in scores from T1 to T2 in the eyes of both the principal and the teachers. Two teachers indicated during the interviews that the mere completion of the training course made them feel 'less supported to reform their practices' than during the course. Somewhat similarly, the school principal stated that 'sustaining a collective sense of urgency among the staff to continue their reform efforts after training would be difficult, but important.'

Lastly, while School 2 showed most improvement in teachers' development compared to the other two schools, the school leadership scores of School 2 appeared poorest among the group. This result thus further supports our previous finding at the treatment level that teachers' attitude and behavior development seems to have operated largely independent of school principals' leadership.

Results of research phase 2

For phase two, we examined the effects of the same intervention on the delayed treatment group (Figure 4.1, boxes *B1* and *B2*). In addition, we investigated the degree to which the first treatment group had managed to sustain its (improved) inquiry teaching practice (Figure 4.1, box *C*).

Analyses at the treatment level

Teachers' attitude development. The effects of the training course on the attitudes of the delayed treatment group were assessed by a 2 (T2 vs. T3) x 4 (*Self-Efficacy* vs. *Context Dependency* vs. *Creative Lesson Design* vs. *Prescribed Lesson Design*) repeated measures MANOVA with time as a within-subjects factor and the four attitude components as the dependent variables. Wilks' statistic revealed no

significant overall interaction effect of time and the four attitude components, Λ = .85, $F(3, 38)$ = 2.11, p = .12, η^2 = .15. However, post-hoc univariate analyses revealed a small but statistically significant effect of time for the *Context Dependency* component, $F(1, 38)$ = 4.12, p = .05, η^2 = .10, which indicates that participants felt less dependent on context factors (MD = −.16) after training (M = 2.63; SD = .95). Notably, Table 4.2 shows that the *Self-Efficacy, Creative Lesson Design,* and *Prescribed Lesson design* scores were already moderate to good at T2.

In line with Figure 4.2 (box *B1*), we also conducted a longitudinal comparison of the pretest scores of the delayed treatment group at T1 with their posttest scores at T3. MANOVA results indicated a large and significant overall interaction effect of time and the four attitude components by Wilks' statistic, Λ = .77, $F(3, 36)$ = 3.54, p = .02, η^2 = .23. Post-hoc univariate analyses revealed a moderate and statistically significant increase in the group's *Self-Efficacy* mean score (MD = .21), $F(1, 38)$ = 4.00, p = .05, η^2 = .10, and decrease in *Context Dependency* mean score (MD = −.29), $F(1, 38)$ = 6.20, p = .02, η^2 = .14. We observed no statistically significant changes in the group's *Creative Lesson Design* and *Prescribed Lesson Design* mean scores during this time period. This result implies that the *Self-Efficacy* and *Context Dependency* mean scores of the delayed treatment group had already positively (but not significantly) developed during phase 1 (T1 vs T2). In our view, these positive score trends during phase 1 might be caused by participants' exposure to the attitude pretest (i.e., test effect) and their awareness of participating in the intervention (i.e., delayed treatment effect), which may have implicitly stimulated positive reflections about their attitudes towards inquiry teaching.

We also found that the mean attitude scores at T1, T2, and T3 of the upper-grade teachers of the delayed treatment group were largely similar to the mean attitude scores of their teams at these times (mean differences smaller than .5 across all attitude components).

Teachers' behavior development

Use of inquiry teaching. Figure 4.2 reveals that, before the second intervention (T2), teachers' *Inquiry Teaching Frequency* (M = 3.36; SD = 1.12) and *Inquiry Teaching Duration* (M = 2.00; SD = .00) mean scores of the delayed treatment group were moderate. Teachers' pupils rated their teachers' behavior by very similar score patterns. After training (T3), we found large and significant improvement in teachers' *Inquiry Teaching Frequency* (Z = −2.50; p = .01) and *Inquiry Teaching Duration* (Z = −3.32; p < .01) mean scores. Teachers used inquiry on a 'daily' basis (M = 4.75; SD = .45) and spend 'half a day' on inquiry at these times (M = 3.00; SD = .00). Their

pupils perceived largely similar improvements during this time. These results indicate that the second intervention also proved successful in increasing teachers' use of inquiry teaching. Teachers' and pupils' personal remarks during the interviews corroborated these results with lesson examples that were very similar to the kinds of examples provided by the first treatment group during phase 1 (T2).

Inquiry teaching quality. Figure 4.2 shows that, after training (T3), teachers' *Cross-Curricular Inquiry Focus* (Z = −2.60; *p* < .01) mean score significantly improved to an excellent score level (*M* = 4.25; *SD* = .45) (all teachers excluded the use of inquiry for teaching mathematics and literacy). Teachers' pupils mean score showed a similar score pattern. No statistically significant changes were found in teachers' *Stimulation of Pupils' Attitudes towards Curiosity* (*MD* = .10), *Stimulation of Pupils' Growth Mindset* (*MD* = .20), and *Stimulation of Pupils' Higher-Order Thinking* (*MD* = .21) mean scores in the eyes of the teachers, despite the apparent positive score trends. Similar to the remarks of the first treatment group, most teachers of the delayed treatment group expressed their struggle with guiding pupils' higher-order thinking during ongoing lesson activities. Several teachers also recommended the use of follow-up training to develop further expertise in this regard. Figure 4.2 shows that pupils' mean scores showed similar score trends, but that their scores were consistently lower than teachers' mean scores across all components, as indicated by Mann-Whitney U test results (all *p*-values < .05). Teachers' *Evaluation of Inquiry Practice* mean scores did improve significantly during this time (Z = −2.07; *p* = .04), but only reached a moderate posttest score level (*M* = 3.00; *SD* = .38). These results indicate that the second intervention affected limited positive changes in teachers' inquiry teaching quality. However, as we will show at the school leadership level, we found several school-specific factors that explain this result.

School principals' behavior development. Figure 4.3 reveals that, prior to the intervention (T2), the school principals of the delayed treatment group perceived their *School Improvement Goals* (*M* = 3.67; *SD* = .57), *School Improvement Priority* (*M* = 4.67; *SD* = .57) and *School Improvement Motivation* (*M* = 4.67; *SD* = .57) for inquiry to be good to excellent, while they perceived their *School Improvement Evaluation* (*M* = 3.33; *SD* = .33), *Inquiry Practice Support* (*M* = 3.55; *SD* = .19), and *Inquiry Practice Adaptation* (*M* = 3.22; *SD* = .19) to be generally moderate. Their teachers rated their school principals' behavior consistently (and substantially) less positive than the school principals perceived their own behavior to be, as indicated by statistically significant results of Mann-Whitney U tests (all *p*-values < .01).

After training (T3), we only noticed a significant decline in school principals' perceptions of *School Improvement Motivation.* We note that their mean score was already exceptionally high to begin with (T2) and thereby posed a natural risk for decline. Two school leaders mentioned in the interviews to experience a 'general drop' in motivation to continue their practice reform now that their team had completed the training course and they had to motivate their school development on their own. Teachers' mean score and remarks during the interview did not reflect this decline. In general, school principals' and teachers' comments reflected similar leadership as mentioned by respondents of the first treatment group (e.g., updated school policy documents, occasional staff meetings about the advancement of the implementation of inquiry teaching).

Lastly, in the eyes of the teachers, school principals' *School Improvement Priority* (*MD* = − .69), Z = −2.85; *p* < .01, and *Inquiry Practice Adaptation* (*MD* = − .17), Z = −2.33; *p* = .02, had significantly declined over time. This result was supported by teachers' remarks during the interviews, but these remarks appeared to be school-specific. We will thus describe these in the next section as part of our analyses at the school level.

Analyses at the school level

   Teachers' attitude development. Mann-Whitney U test results revealed that, at T2, the mean scores of School 6 on the *Context Dependency* and *Prescribed Lesson Design* components were significantly lower than the mean scores of School 4 and 5 on these same components (all *p*-values < .04). This means that, prior to the intervention, the teachers of School 6 generally felt less dependent on context factors and less confined to prescribed lessons to teach through inquiry. At T3, Mann-Whitney U test results indicated that the mean scores of the three delayed treatment schools were statistically similar for the *Self-Efficacy* and *Prescribed Lesson Design* components, but that School 6's mean scores on *Context Dependency* and *Creative Lesson Design* were significantly better than of School 4 and 5 (all *p*-values < .05).

   Wilcoxon Signed-Rank test results revealed no statistically significant changes in schools' mean attitude pretest (T2) and posttest (T3) scores during phase two. However, as part of our longitudinal investigation (T1 vs. T3), Wilcoxon Signed-Rank test results did reveal a statistically significant decrease (*MD* = −.32) in School 5's *Prescribed Lesson Design* mean score (Z = −2.08; *p* = .04). We note that the teacher teams of School 4 and 6 were smaller than the teacher team of School 5 and thus

provided less statistical power to indicate significant changes in scores, despite their largely similar positive mean score trends for this component.

According to the school principal and the teachers of School 6, the team already worked with new approaches to teaching and learning in their school (e.g., the use of individual lessons plans per child, cross-curricular education), which had made their team already familiar with some features of inquiry-oriented teaching. We believe that this prior experience had partially fostered School 6's attitudes towards inquiry teaching prior to the program. In fact, the school principal and the teachers expressed their discontent about the attitude-focused training course to the first trainer and during the interviews at T2, as in their eyes, the contents of this training turned out to provide little added value to the professional development of the team.

Teachers' behavior development

Use of inquiry teaching. In particular, we note that Figure 4.2 indicates a maximum scale score (5) on the *Inquiry Teaching Frequency* component by the teachers of School 6 prior to (T2) and at the end of the program (T3). Their pupils agreed that their teachers used inquiry on a daily basis at both times. Thus, in line with our previous findings, these results further show that School 6 seemed more familiar with inquiry-oriented teaching than Schools 4 and 5 at the onset of the program.

Inquiry teaching quality. Furthermore, Figure 4.2 reveals that teachers' *Cross-Curricular Inquiry Focus* mean score of School 6 was also at the scale maximum before (T2) and after the program (T3) in contrast to School 4 and 5. The teachers explained that their school already used lesson methods that adhered to thematic, cross-curricular approaches to teaching and learning, but in their eyes, still lacked emphasis on stimulating pupils' inquiry.

Despite of School 6's high scores on *Inquiry Teaching Frequency* and *Cross-Curricular Inquiry Focus* compared to School 4 and 5, we found no substantial differences in the mean scores on most of the remaining components of teachers' inquiry teaching quality between the three delayed treatment schools (overall moderate scores). In the case of *Inquiry Teaching Evaluation*, we did find noticeable improvement in teachers' mean score of School 6 (i.e., reaching a moderate score level). Also, in the case of teachers' *Stimulation of Pupils' Higher-Order Thinking*, we note substantial improvement in pupils' perception scores at School 5 (i.e., reaching a moderate mean score). However, this improvement was reflected less prominently in teachers' own perception scores. This result might indicate that the pupils became

more aware of teachers' (moderate) stimulation of their higher-order thinking over time.

The teachers of School 5 explained their moderate and stagnant inquiry teaching scores (T3) due to an overall 'lack of urgency' in their team to reform their practices. According to the teachers and their school principal, this was largely caused by their 'rough start of the program', since the team collectively decided to stop the attitude course due to their concerns about the 'personal style' of the first trainer. The team of School 6 also expressed their discontent about the attitude course, but instead felt that the course did not met their professionalization needs, which is supported by our attitude data. As such, the teachers of School 6 stated that they felt less committed to fully engage in the remainder of the training course. At School 4, our interview data provided less evidence to explain teachers' lack of inquiry teaching development.

School principals' behavior development. Figure 4.2 shows no substantial differences in the mean scores of the three delayed treatment schools on most components of school principals' leadership behavior. In the case of *School Improvement Goals*, we note that the school principal of School 6 rated *School Improvement Goals* significantly less positive than the school principals at School 4 and 5. In contrast to School 4 and 5, the school principal of School 6 explained that she perceived herself unable to set meaningful school improvement goals for inquiry-based teaching practice for her school on the basis of the training received, since in her eyes, the training turned out to provide little added value to the professional development of the teacher team. Teachers' mean scores of School 6 on *School Improvement Goals* were moderate as well and, in the eyes of these teachers, were justified for the same reason. Despite these personal remarks, however, we note that the mean pretest and posttest scores of the teachers at School 6 on inquiry teaching quality and evaluation were only moderate at best and thus allowed significant room for improvement.

Lastly, we note that the teachers at each delayed treatment school provided different explanations for their perceived decline in school principals' *School Improvement Priority*. The teachers of School 4 and 5 described competing education reform initiatives in their schools during phase 2 that demanded their attention (e.g., new lesson methods). At School 6, the teachers perceived that their school principal felt less convinced about the relevance of the training program. During most course meetings, both trainers independently noted School 6's general lack of engagement in completing the training exercises (e.g., little group discussion and brainstorming, unfinished homework assignments).

Comparison of the posttest scores of the first and the delayed Treatment

Next, we examined to what extent both interventions achieved similar posttest scores among the six treatment schools immediately after course completion (Figure 4.1, box *B2*).

Analyses at the treatment level. No statistically significant changes were found between the mean attitude posttest scores of the first treatment group (T2) and the delayed treatment group (T3). Figure 4.2 and 4.3 indicate that the first and the delayed treatment group had attained largely similar posttest scores across most of our inquiry teaching behavior and school leadership measures. Mann-Whitney U test results revealed that teachers' and pupils' posttest scores on the *Stimulation of Pupils' Attitudes towards Curiosity, Stimulation of Pupils' Growth Mindset*, and *Evaluation of Inquiry Practice* components of the first treatment group (T2) were significantly more positive than teachers' and pupils' posttest scores (T3) of the delayed treatment group on these components (all *p*-values < .05). Figure 4.2 shows that the intervention improved teachers' inquiry teaching quality only in the case of the first treatment group, but not the delayed treatment group. These above results indicate that, although we intended the intervention to affect similar improvements for all participants, the intervention proved significantly more successful in fostering inquiry teaching quality for the first treatment group than for the delayed treatment group. As described, we believe that this result can be explained by several incidental factors that hindered the school improvement of the delayed treatment group. School leadership at both treatment groups was characterized by largely similar posttest across all components.

Analyses at the school level. At the individual school level, Mann-Whitney U test results revealed significant differences in teachers' *Context Dependency* pretest and posttest mean scores between School 6 and the other five treatment schools (all *p*-values < .05). In this case, School 6's proved to perceive itself least dependent on context factors for teaching through inquiry. In addition, teachers' and pupils' *Cross-Curricular Inquiry Focus* mean pretest scores of School 6 were highest among all treatment schools (all *p*-values < .05). These results thus reveal that School 6 proved most experienced among the six treatment schools at the onset of the intervention.

In the case of school principals' leadership, we note that School 3 outperformed most other treatment schools on the basis of teachers' perception mean scores, particularly in the cases of *School Improvement Goals*, *School Improvement Evaluation*, *Inquiry Teaching Support*, and *School Improvement Motivation* (i.e.,

moderate to good scores). This score pattern seems less apparent on the basis of school principals' perception scores on these components.

Longitudinal effects of the first treatment

Lastly, we examined the degree to which the first treatment group managed to sustain its improved behavior scores after having completed the training course (T2) to the end of phase two (T1 vs. T2 vs. T3) (see Figure 4.1, box *C*).

Analyses at the treatment level

Teachers' behavior development. Wilcoxon Signed-Rank test results revealed that the first treatment group had successfully sustained teachers' *Inquiry Teaching Frequency* and *Inquiry Teaching Duration* during phase two in the eyes of the teachers and their pupils, as indicated by non-significant mean score differences (T2 vs. T3).

In addition, we found that the pupils of the first treatment group perceived significant improvement in teachers' *Cross-Curricular Inquiry Focus* from T2 to T3, $Z = -2.50$, $p < .01$, during this time (i.e., most pupils indicated that their teachers increasingly used inquiry for teaching school topics related to 'social studies'). However, this mean score increase was not found on the basis of teachers' own mean score on this component. In our view, this finding suggests that the pupils may have simply become more knowledgeable and/or aware during the second phase of some of the features of teachers' cross-curricular focus of inquiry teaching.

Furthermore, teachers' high mean scores on *Stimulation of Pupils' Attitudes Towards Curiosity, Stimulation of Pupils' Growth Mindset,* and *Evaluation of Inquiry Practice* were sustained to the end of the next school year (T3). Pupils' score patterns seemed to be very much alike, except in the case of teachers' *Stimulation of Pupils' Attitudes Towards Curiosity,* which showed significant improvement ($MD = .46$) from T2 to T3, $Z = -2.72$, $p < .01$. Because teachers' mean scores on this component had remained unchanged during this time, pupils' score increase might again indicate that pupils mainly developed in their understanding or awareness of teachers' behavior aimed at stimulation their attitudes towards curiosity.

The above results were corroborated by personal remarks of the teachers and pupils during the interviews (T3). All the teachers emphasized their continued efforts to teach school subject matter through inquiry during the second research phase and indicated that inquiry-based pedagogy felt 'more and more part of their school's usual day-to-day business' or 'much more like their own ordinary perspective on what education should be like'. Most teachers explained their continued work by their

increased feelings of confidence and independency to teach school subject matter through inquiry. Also, more than half of the teachers stated that, in their eyes, teaching through inquiry had become increasingly enjoyable during the second phase. For example, one teacher perceived 'the positive effects of inquiry on pupils' learning motivation and involvement during class' and thereby felt motivated to keep using inquiry. Many teachers also claimed that their inquiry teaching revealed 'hidden academic talents and motivations of [their] pupils that [they] had not seen before during traditional, teacher-directed teaching', which made them increasingly convinced about the educational relevance and please of teaching school subject matter through inquiry. These results suggest that teachers' continued inquiry teaching behavior was mostly driven by their (increasingly) positive attitudes towards inquiry teaching.

Notably, our analysis also revealed significant improvement in teachers' *Stimulation of Pupils' Higher-order Thinking* of the first treatment group during the second phase in the eyes of the teachers (*MD* = .30; $Z = -2.33$; $p = .02$) and their pupils (*MD* = .33; $Z = -2.05$; $p = .04$). Several teachers explained that their ongoing familiarization with inquiry teaching had made inquiry teaching 'less mentally taxing' for them and, consequently, allowed them to better focus on learning to encourage pupils' higher-order thinking. Their pupils indicated that teachers had become 'stricter about [them] needing to come up with solutions, questions or ideas that were more original and surprising.' Other pupils described that their teachers involved them more often in rephrasing typical study assignments into 'more interesting and exciting assignments for [them] to work on'. These results indicate that, only after having completed the training course, the first treatment group had started to overcome their initial difficulties with stimulating pupils' higher-order thinking.

School principals' behavior development. During phase two, the mean perception scores of the school principals and the teachers of the first treatment group had remained largely stable for all components of school principals' leadership. However, we note that this observed 'stability' may perhaps not amount to much accomplishment, as the mean posttest scores of the first treatment (T3) appear very much equal to their scores before the start of the program (T1). In fact, none of the school principals mentioned anything during the interviews about 'deliberate efforts' aimed at sustaining their school leadership during the intervention.

We did observe a significant decrease (*MD* = − .24) in teachers' *School Improvement Goals* mean score ($Z = -2.27$, $p = .02$) of the first treatment group during phase two (T2 vs. T3). This development was corroborated by comments of

several teachers during the interviews (T3). In teachers' eyes, their school principals had become less clear about formulating new school improvement goals to maintain and optimize their inquiry-based teaching practice now that they had completed the training course. Few teachers mentioned to desire revisions of pupils' school performance cards to also include descriptions of pupils' inquiry performance (e.g., curiosity, higher-order thinking, growth mindset). This issue was not mentioned by any of the school principals during the interviews at T3.

Analyses at the school level

Teachers' behavioral development. During phase two, we observed no substantial mean score differences between the first treatment schools on most of the components of teachers' inquiry teaching behavior. Figure 4.2 shows increased positive trends in teachers' *Stimulation of Pupils' Attitudes Towards Curiosity* mean scores of the teachers and the pupils of School 2 compared to teachers' and pupils' mean scores of School 1 and 3 on this component during this time. Figure 4.2 also indicates a significant decline in the mean score of the pupils of School 3 on teachers' *Stimulation of Pupils' Growth Mindset.* Lastly, we also observed substantial improvement in teachers' *Stimulation of Pupils' Higher-Order Thinking* scores in the eyes of the teachers and the pupils of School 2. In all of the above cases, however, our interview data did not provide substantive explanations for these school-specific score patterns.

School principals' behavior development. Figure 4.3 reveals that during phase two, across most school leadership components, School 3 continued to outperform School 1 and 2 on the basis of school principals' and teachers' perception scores. No significant school-specific changes in school leadership scores were found for most components during this time. In the case of *School Improvement Motivation*, Figure 4.3 did show a continued decline in teachers' mean score of School 3 (T1 vs. T2 vs. T3). Three of these teachers claimed during the interview (T3) that their school principal came over as 'less clear', 'less concrete' or 'less convincing' about what work was needed to continue their reform efforts now that the training was complete. However, we note that teachers' posttest scores (T3) of School 3 on this component were still good, as shown in Figure 4.3.

Discussion

Main findings

Teachers' attitude development. The present results revealed that teachers' self-efficacy and context dependency mean pretest scores (T1) of both treatment groups were equally moderate. After the program, in line with our expectations, result indicated that teachers' self-efficacy mean scores had significantly increased while their context dependency mean scores had significantly decreased in the case of the first treatment group (T1 to T2) and the delayed treatment group (T2 to T3) (H1 supported).

In addition, the results indicated that the professionalization treatment positively affected teachers' creative lesson design belief mean scores in both treatment groups. However, no clear significant changes were observed in participants' prescribed lesson design mean scores during the program. Thus, H2 was partially supported. We note that, prior to the program (T1), participants' lesson design belief scores were already quite positive to begin with. These positive scores thus left little room for further improvement during the intervention and did not help us differentiate much between the participating school teams in both treatment groups.

The above findings were largely corroborated by our qualitative interview data. Many teachers indicated to increasingly 'enjoy' teaching through inquiry, as in their eyes, they became more aware of the positive effects of inquiry-based teaching on their pupils' academic performance (e.g., greater persistence, new talents). In our view, this result suggests that teachers' attitude development may well have continued after training. Therefore, we regret not having measured attitude among the first treatment group at the end of phase two (T3) in order to quantitatively verify this reported development among the teachers.

Lastly, most school leaders and teachers of both treatment groups stated during the interviews to greatly value the attitude training as part the intervention. In particular, respondents valued that their whole teams participated in the training, because they perceived it stimulated feelings of school-wide commitment to bringing about practice reform. These claims are in line with general research on school improvement (e.g., Desimone, 2009).

Teachers' behavior development. The above positive changes in teachers' attitudes coincided with a large increase in teachers' inquiry teaching frequency and duration in both treatment groups (H3 supported). This result is very encouraging as

well, because one of the main aims of the program was to increase teachers' use of inquiry-based teaching. Once again, this finding supports findings by Van Aalderen-Smeets and Walma van der Molen (2015) and supports the proposition that positive attitude development positively effects teaching behavior development.

However, improvements in teachers' inquiry teaching quality were only observed for the first treatment group (H4 partially supported). From T1 to T2, teachers' *Cross-Curricular Inquiry Focus*, *Stimulation of Pupils' Attitudes Towards Curiosity*, *Stimulation of Pupils' Growth Mindset,* and *Evaluation of Inquiry Practice* mean scores significantly increased. These improvements were corroborated by teachers' personal statements during the interviews. Notably, however, many teachers stated to teach 'numeracy and literacy' through direct instruction and used inquiry mainly for teaching history, geography, science and social studies. Some teachers desired specialized training for teaching these traditional subject domains through inquiry.

However, on average, the first treatment group showed marginal improvements in their stimulation of pupils' higher-order thinking. In the eyes of many teachers, this appeared to be one of the most difficult and challenging tasks of their inquiry teaching practices. At the same time, we note that the first treatment group showed modest improvement on this measure during the *second* research phase (T2 to T3), which suggests that teachers might have needed more time to further develop the skill. Teachers' remarks during the interviews suggest that, indeed, they seemed better able to focus on the more challenging aspects of their inquiry teaching practices when other important inquiry teaching tasks were mastered.

Nevertheless, the first treatment group successfully sustained their improved use and quality of inquiry teaching practice to the end of the next school year (RQ1). Many teachers of this group stated that their sustained practices were mainly driven by their collective and positive 'convictions' about 'feeling in control' of their inquiry teaching practices. This result highlights the value of attitude-focused teacher professionalization also for realizing *sustained* school improvement (see also Thurlings et al., 2015).

In the case of the delayed treatment group, however, the same (delayed) intervention did not prove to be as successful in improving teachers' inquiry teaching quality of the delayed treatment group compared to the first treatment group (T2 to T3). We believe that the group's overall lack of teaching quality development can be largely explained by school-specific factors that hindered school improvement (i.e., mismatch of training needs for School 6, School 5 being discontent about the personal

style of one of the trainers, new lesson methods at School 4 and 5). This explanation was provided by the school principals and teachers too. Thus, we feel safe to say that the positive effects found for the first treatment group in this study can be attributed to the intervention. Nonetheless, these findings highlight the value of examining the efficacy of school improvement interventions among other primary school teams at other times, since our results indicate the influence of 'incidental' factors on the success of such interventions.

School principal's behavior development. Lastly, we set out to examine the degree to which school leadership on teachers' adoption of inquiry-based pedagogy would develop among the treatment schools during the intervention (RQ2). In addition, we wanted to know what school leadership factors might matter in this regard (RQ3). Overall, most teachers of both treatment groups stated that their school principal showed persistently poor performance in communicating clear policy on practice reform and in evaluating these goals during staff meetings or classroom visits. Many teachers also complained about the lack of support and guidance they were provided by their school principal on their practice reform during the program. Although school principals' perceptions about these matters were more positive than teachers' perceptions, several school principals did indicate to desire specialized leadership training to learn to better support teachers' adoption of inquiry-based teaching practice. In our view, these results suggest that improved school leadership for the implementation of inquiry-based teaching is thus not a natural "by-product" of extensive and whole-team inquiry-focused teacher professional development. A more explicit focus on the development of school leadership on inquiry teaching practice reform could thus be fruitful for enhancing such leadership.

Given the above, however, we note that the teachers largely improved their inquiry teaching behavior and teaching quality *despite* largely poor school leadership. The present results thus support the predominant notion in the school improvement literature that the effectiveness of education reform efforts in schools can be explained more by what teachers do and think than by factors at the school leadership level (Muijs et al., 2014). Therefore, the question rises whether more explicit attention for developing school leadership in the context of the current study would actually make a significant difference in teachers' inquiry teaching development.

Because our sample consisted of only a small number of primary schools, of which the majority showed stable and largely similar leadership over time, we were unable to uncover significant interactions between our teacher variables and school leadership variables at the individual school level. We were thus unable to adequately

answer our third research question. However, we did observe small, but statistically significant differences in teachers' mean attitude scores (e.g., School 6 showed significantly better mean attitude scores compared to the other schools), teachers' behavior scores (e.g., School 2 showed significantly better behavior scores compared to the other schools) and school principals' leadership quality (e.g., School 3 showed significantly better mean leadership scores compared to the other schools) among the treatment schools on several of our dependent variables. This indicates that the degree to which the school teams adopted inquiry-based pedagogy differed somewhat during the program. As others have noted as well, such differentiation is to be expected, since schools naturally differ in terms of their teacher demographics, constellations, professional routines, and organization (for details about this matter, please see the Concerns-Based Adoption Model; Horsely & Loucks, 1998; Loucks-Horsley, 2001).

Limitations and future research
The above-described results of the present study are promising, but should also be interpreted in light of several methodological limitations. First, we are careful with generalizing our results. The six primary schools that participated in the study were largely self-selected based on their interests in the goals of the current school improvement program. In addition, all schools were required to pay for the professional training. While we believe that this ensured high levels of commitment among the participating school teams, which was believed necessary for successful practice reform (see also Desimone, 2009), our current findings might thus apply only to *committed* primary school teams. Future studies should therefore investigate the efficacy of similar inquiry-focused school improvement programs among truly naïve primary school teams and without requiring schools to pay for enrollment. Such research seems highly relevant, since current international education policy documents increasingly promote nationwide implementations of inquiry-based pedagogy in primary education (e.g., OECD, 2015), which would likely involve primary school teams that do not (yet) perceive the relevance and urgency of adopting such pedagogy.

Extending on the above point, our delayed treatment control group design did not provide a true naive control group comparison during the intervention. Although no significant differences were found in the mean attitude pretest scores (T1) between the two treatment groups and our baseline control group, results did reveal slight (but statistically non-significant) improvements in the attitudes of the delayed treatment

schools during the first research phase. This suggests the presence of testing effects and/or delayed treatment effects. These effects may perhaps have been more widespread and structural throughout the intervention of the delayed treatment group than we were able to observe and thereby may have confused the current program effects. However, recruiting school teams for a naïve control condition in any longitudinal school development research project is difficult, because school teams that do not perceive the relevance of such a project are not likely to commit themselves to multiple and extensive measurements without receiving the benefit of training. Therefore, our delayed treatment control group design proved useful.

We also note that we did not compare the effects of the present intervention with other professional development programs. Thus, one could argue that it is not surprising that training has more effect than no training. For example, it would have been valuable to compare the effects of our current program to a group of schools that received traditional inquiry-focused professionalization (e.g., implementation of prescribed inquiry lessons or attitude training only). This way, future studies could investigate the contribution of particular features of inquiry-focused school improvement programs to teachers' potential inquiry teaching development.

Since the aim of our intervention was to positively impact the inquiry teaching practices of *all* teachers in the school organization, we are aware that our observed changes in teachers' inquiry teaching behavior were limited to the upper-grade teachers. As described in the Method section, we decided this for practical reasons. Although we assume that our findings at the upper-grade levels can be generalized to reflect similar changes in the inquiry teaching behavior of the entire teacher teams, this assumption needs to be further tested. We do note that the mean attitude scores of the upper-grade teachers in the current sample were largely similar to the mean scores of their school teams, which suggests that the upper-grade teachers showed similarity with their school teams. Nevertheless, we recommend that future studies assess the impact of similar interventions on the practice reform of lower-grade teachers as well.

In this study, we assessed (changes in) inquiry teaching behavior and school leadership only by means of perception data. We are aware that post-hoc interviews make respondents prone to providing socially desirable responses. However, because the internal consistency of our obtained perception data proved to be sufficient, the perception data of the related respondent groups showed similar patterns, and respondents provided fitting illustrations of their perception scores, we believe that our perception data provided reliable insight into the realized school improvement.

Across most of our measures, respondents rated their own behavior more positively than their peers rated this behavior to be, which is in line with results found by others (e.g., Anderson et al., 2010). However, we note that our interview instrument was newly developed and will thus require further optimization and validation. For example, we recommend research into the concurrent validity of our perception scales, such as by comparing teachers' and pupils' perceptions about teachers' inquiry teaching behavior with classroom observations by external raters to see whether these different measures converge to similar conclusions. The current interview format could also be translated into a self-report questionnaire format, allowing researchers to more efficiently collect data among individual respondents and, in addition, gain insight into the variance among respondents' individual perception scores to examine respondents' agreeability.

Lastly, we did not find significant interactions between school leadership and teachers' inquiry teaching behavior. As discussed, the majority of the participating schools differed only marginally in their leadership scores, which complicated our analyses (e.g., comparing teachers' behavior development in schools with high versus low leadership scores). Future research with more large-scale data collection may likely result in a wider distribution of scores among participating schools, which might provide better insight into the possible relations between our independent and dependent variables. In addition, such large-scale research may also improve the statistical power of other analyses and allow for more advanced (hierarchical) statistical testing.

Implications

Notwithstanding the above limitations, we believe that the positive effects found in this study can be credited to the efficacy of the school improvement program. In our view, the present findings have important implications for research and practice.

Fostering teachers' attitudes towards inquiry teaching. In sum, the present results support that the attitude-focused training by Van Aalderen-Smeets and Walma van der Molen (2015) indeed successfully improved teachers' attitudes towards inquiry teaching. We highlight that school teams who are highly committed to adopt inquiry-focused teaching practice, do not necessarily possess positive attitudes towards inquiry teaching to begin with. As shown, teachers' self-efficacy and context dependency mean pretest scores regarding inquiry teaching were moderate. This finding is in line with Van Aalderen-Smeets and Walma van der Molen (2012, 2015) and attitude training should thus be regarded a vital and explicit starting point

of inquiry-focused teacher professionalization. This seems to be true even in the case of primary school teams that already perceive the relevance and urgency of such professionalization.

However, on average, teachers' creative lesson design perception pretest scores were already quite positive. This result is somewhat surprising, since research indicates that primary teachers generally feel inclined to teach inquiry through the use of prescribed inquiry lessons and therefore require encouragement to foster pupils' inquiry creatively on their own (e.g., Jones & Eick, 2007). In The Netherlands, however, the government generally decides what primary schools should teach, but schools are free to decide *how* they will teach subject matter. This fact may explain teachers' naturally high creative lesson design and low prescribed lesson design pretest scores in our sample. Alternatively, teachers' high creative lesson design scores may simply be the result of self-selection bias: *because* the school teams already held positive perceptions about creative inquiry-focused lesson design, they decided to enroll into the current program (to learn *how* to do so). In any case, teachers' positive perceptions of lesson design thus left little need for further improvement. The current teacher training course could thus be shortened by spending less time on the positive development of primary teachers' creative lesson design perceptions. A shortened professionalization program would likely appeal to prospective primary schools, as most primary school teams generally have limited time and resources available for training and practice reform.

Similarly, we propose that the attitude-focused part of the training course could perhaps be further shortened by focusing on the positive development of teachers' self-efficacy and context dependency beliefs alone (excluding a focus on teachers' perceived relevance, difficulty, and enjoyment of inquiry teaching). Findings by Van Aalderen-Smeets & Walma van der Molen (2015) indicate that teachers' self-efficacy and context dependency beliefs function as particularly salient drivers of their potential inquiry teaching behavior. Indeed, in the current study, many teachers commented that their (collective) self-efficacy and context dependency beliefs mainly drove their inquiry teaching behavior development. Moreover, in line with the findings of Van Aalderen-Smeets & Walma van der Molen (2015), teachers' pretest scores on these measures were moderate, which underscores the relevance of professionalization aimed at the positive development of these variables. Attitude training focused on teachers' self-efficacy and context dependency beliefs might prove to be enough to significantly foster teachers' willingness to adopt inquiry teaching practice.

To stimulate teachers' attitude development further, we recommend that teacher-trainers encourage teachers to actively discuss their current and planned implementation of inquiry teaching throughout the course program. As observed at School 2, participants' overall lack of attitude development during the course corresponded with their overall lack of 'collective discussions' about implementation. Conversely, at the other schools, many teachers illustrated the ways such discussions challenged and fostered their (shared) beliefs about reforming their practices. Therefore, we emphasize that active discussions among all school staff about the implementation of the course content should be made an integral and explicit component of professionalization treatment. We also note that this finding underscores the value of *school-wide* teacher professionalization, as this provides teachers more opportunity to discuss their (beliefs about) practice reform with peers.

Lastly, the current findings suggest that teachers' attitude development may well have continued after training and may have involved additional components of their attitudes, such as teachers' perceptions of 'enjoying' to teach through inquiry. Notably, the DAS questionnaire (Van Aalderen-Smeets & Walma van der Molen, 2013) includes the measurement of this attitude component, among other components. In the recent intervention study by Van Aalderen-Smeets and Walma van der Molen (2015), however, no significant improvement was found in participants' scores on this component immediately after training. The results of the present study may suggest that teachers need more time to fully mature in their attitudes towards inquiry teaching and that the development of teachers' attitudes towards inquiry teaching does not occur simultaneously across all components, but instead might occur on a per-component basis in different phases. This finding underscores the importance of longitudinal measurements of teachers' professional development.

Teachers' inquiry teaching skills. As discussed, most teachers perceived the stimulation of pupils' higher-order thinking to be particularly difficult. According to Miri, David and Uri (2007), teachers need to possess higher-order thinking skills themselves to be able to 'creatively' respond to pupils' inquisitive ideas and questions during lesson activities. Clearly, not all primary teachers necessarily possess these higher-order thinking skills. Our findings suggest that primary teachers may thus benefit from follow-up coaching by expert trainers or experienced peers to master this complex inquiry teaching skill, on the job, over a longer period of time (see also Miri et al., 2007). Such coaching should involve the development of teachers' own higher-order thinking skills, such as by conducting inquiry themselves, but also further

familiarization with strategies to stimulate pupils' higher-order thinking across different classroom scenarios, such as by means of asking pupils different types of questions, fostering group discussion, or extending on pupils' inquisitive ideas during lesson activities by proposing design experiments.

Throughout the program, many teachers also stated to teach 'numeracy and literacy' by means of direct instruction, while they used inquiry mainly for teaching social and environment studies. In their view, direct instruction proved more effective to achieve their numeracy and literacy lesson objectives. However, this does not necessarily imply that numeracy and literacy education cannot be meaningfully incorporated into inquiry-based lesson activities. In fact, inquiry-based learning can provide a particularly powerful context for pupils' language development, since language serves pupils to describe, predict, and explain findings derived from their inquiries (Stoddart, Pinal, Latzke, & Canaday, 2002). Similarly, mathematics may serve pupils' inquiry as well (Makar & Fielding-Wells, 2017). Such integration may even be reciprocally beneficial: as pupils conduct inquiry, they discover new subject matter, which demands using (new) language and mathematics to make sense of their discoveries, which increases their conceptual understanding of the subject matter at hand, which may foster further inquiry into the subject matter, etcetera. To this end, teachers should be trained to perceive the opportunity and educational relevance of such integrative lesson activities. For example, as part of their professionalization, teachers should be given more concrete lesson examples of what such integrative lesson plans may look like and practice devising such lesson plans themselves under guidance of trainers.

Leadership on inquiry teaching practice reform. Our findings show that school principals' leadership development on inquiry-focused teaching practice reform is not implicated through their participation in inquiry-focused teacher professionalization. Throughout the program, many teachers indicated to appreciate the positive cultural norms set by their school principal for school improvement, but felt that their school principals fell short in providing clear expectations, support and guidance concerning their own teaching practice reform. This shortcoming was perceived by several school principals too, who therefore indicated a desire for specialized leadership training. Put differently, this result implies that school principals' commitment to adopting inquiry-based pedagogy in their school thus provided little bearing on the quality of policy, support, and guidance they provided teachers for successful practice reform. This finding is in line with Timperley (2008).

Research by Miller et al. (2016) suggests that school leadership training should at least include opportunities for school principals to collaborate with and learn from other (successful) school principals (Hamilton, Engberg, Steiner, Nelson, & Yuan, 2012). To this end, school district boards should establish professional learning communities of school principals to make such collaborations possible. School principals could exchange best practices and discuss shared challenges to foster inquiry-focused practice reform among their teacher teams. In addition, Grissom and Harrington (2010) indicate that school principals' participation in formal mentoring programs at their own school prove to be more effective to foster leadership than participating in formal (university) courses on leadership (see also Miller et al., 2016). Such mentoring should aim at teaching school principals to regularly probe teachers' experiences with adopting inquiry-based pedagogy, such as by classroom visits, and to set meaningful performance goals and devise strategies at the school level to achieve these goals (see also Murphy & Seashore, 2018). These strategies may involve the re-allocation of teachers' time to allow for more frequent teacher collaboration and classroom experimentation, role modeling to teachers what it means to be 'inquiry-minded' towards school practice reform, or introducing best practices obtained from other schools. In our view, such mentoring of school principals should be integrated into school development programs.

Perception data to assess inquiry-based teaching practice. Lastly, the present findings suggest that pupil perceptions can be efficiently used for evaluating inquiry teaching quality. In the upper grades, pupils were able to distinguish between various behavioral measures that relate to their teachers' inquiry teaching behavior (see also Van der Scheer et al., 2018). Teachers can thus perhaps be trained to gather such pupil perception data themselves and use the results of these evaluations to improve their own inquiry teaching quality where needed. In addition, we propose that school principal, teacher and pupil perception data can be used by teacher-trainers and researchers to better assess the specific professionalization needs of school teams and, subsequently, to develop more targeted professionalization trajectories.

Conclusion

The main results of the current study support our assumption that a combination of extensive attitude-focused and didactical training provides primary school teams the (minimal) preparatory training to adopt inquiry-based teaching practice in their

schools. This finding implies that such implementation in primary education does not necessarily require invasive changes in the school curriculum, school organization, or teaching and lesson materials. The teachers in our study were taught to implement inquiry teaching in small steps in their existing lessons, such as by asking pupils different kinds of questions, evoking discussion by connecting school topics, and encouraging pupils to explore novel knowledge domains.

At the same time, our findings highlight that it is not easy for primary teachers to implement inquiry-based pedagogy on their own. Changing habits is difficult, especially when change requires the mastery of new complex teaching skills. Future research is needed to learn how inquiry-based pedagogy can be further promoted in primary education. To this end, we need more longitudinal quantitative and qualitative research, aimed at evaluating the impact and exploring the processes of implementation, while considering the complexity of schools as organizations, and refraining from overly prescriptive approaches to encouraging and assessing teachers' inquiry-pedagogy. The findings of this study may serve as a starting point for the further development and evaluation of inquiry-focused school development interventions.

Chapter 5

5 Effects of an inquiry-focused school improvement program on the development of pupils' attitudes towards curiosity, their implicit ability and effort beliefs, and goal orientations

Abstract

The present study describes the effects of a two-year inquiry-focused school improvement intervention on pupils' attitudes towards epistemic curiosity, their implicit beliefs about the malleability of their ability, their effort beliefs, and their goal orientation motivations were investigated. Six Dutch primary school teams participated in the study. Quantitative data were collected from the 4th, 5th and 6th grade pupils and examined based on a longitudinal pretest-posttest control group design. A Structural Equation Modeling approach was first used to examine the relationships among pupils' attitude, belief, and motivation scores. In line with attitude and motivation theory, pupils' attitudes towards epistemic curiosity and their implicit ability beliefs predicted their efforts beliefs and goal orientation motivations. In addition, the intervention affected positive changes in pupils' attitude, belief and motivation scores over time. Findings may inform the further development of school development projects aimed at stimulating pupils' inquiry in primary education, namely by fostering pupils' attitudes towards epistemic curiosity and their implicit ability beliefs.

Introduction

Policy documents on twenty-first century learning increasingly promote primary school curricula aimed at engaging pupils in inquiry (Lucas, Claxton, & Spencer, 2013; OECD, 2015). Inquiry has been associated with a wide range of complex learning behavior, such as voicing inquisitive ideas and questions, creatively linking subject matter from different knowledge domains, and seeking alternative answers or solutions to questions or problems (e.g., Engel, 2015; Heywood, Parker, & Jolley, 2012; Osborne, 2014). When pupils learn by inquiry, it is believed that they develop a more meaningful and integrated understanding of school subject matter than by forms of rote learning (Pellegrino & Hilton, 2012).

Despite the educational value of inquiry-based learning, scientific descriptions of school-wide interventions aimed at fostering pupils' inquiry thinking are still scarce and their effects are often not studied on the level of the pupil (Syer, Chichekian, Shore, & Aulls, 2012). Therefore, the present authors developed a two-year school improvement program in which primary school teams were trained to adopt inquiry-based pedagogy. To stimulate pupils' inquiry, teachers learned how to stimulate pupils' (a) attitudes towards epistemic curiosity (i.e., perceived value of inquisitive thinking), (b) implicit ability beliefs (i.e., perceived malleability of their ability), (c) positive effort beliefs (i.e., perceived causality of their effort on achievement), and (d) achievement goal orientations (i.e., perceived goals for achieving in school), and to adapt their daily practices accordingly. Previous research suggests that these variables are particularly salient precursors to pupils' inquiry and can be fostered through school interventions (an overview of this research will be provided later on in this Introduction). Results from a recent experimental study showed that the two-year school improvement program indeed increased teachers' inquiry-based pedagogy (Post & Walma van der Molen, in press). The question remains, however, whether teachers' improved inquiry teaching behavior also affected subsequent improvement in their pupils' attitudinal and motivational beliefs that relate to inquiry.

The present paper aims to answer the above question. To this end, we first examined the *relationships* among pupils' attitudes, beliefs and goal orientations. Based on attitude and motivation theory, we propose that attitudes towards epistemic curiosity (Post & Walma van der Molen, 2018b) and implicit ability beliefs (Dweck, 2000) may predict effort beliefs and achievement goal orientation motivations (e.g., Ajzen, 2001; Ajzen & Fishbein, 1980; Blackwell, Trzesniewski, & Dweck, 2007; Eagly & Chaiken, 1993; Grossnickle, 2016; Tempelaar, Rienties, Giesbers, & Gijselaers,

2015; Van Aalderen-Smeets & Walma van der Molen, 2018a). Specifying these hypothesized causal relations into testable models, as done in this paper, is important for two reasons. First, it may provide empirical evidence for these proposed relations. And second, if attitude towards epistemic curiosity and implicit ability beliefs indeed *affect* effort and motivation, the stimulation of positive attitudes towards epistemic curiosity and implicit ability beliefs may be a valuable supplement to other forms of inquiry-based pedagogy. As far as we know, these hypothesized relationships have not been tested before in the context of an inquiry-focused school intervention.

Second, in the present study, we investigated the effects of the school intervention program on pupils' attitudes towards curiosity, implicit ability beliefs, effort beliefs, and achievement goal orientation motivations. This was done by comparing the changes in pupils' attitude, belief, and motivation scores during the school intervention program to pupils whose teachers had not (yet) participated in the program. Findings provide insight into the malleability of pupils' attitudes, beliefs, and motivations by teachers in a real-life school context and, thereby, into the efficacy of a school-wide implementation of inquiry-based pedagogy in primary schools. This insight may inform the further development of such pedagogy for primary education.

Before providing the theoretical foundations of the present study, we first provide a summary of the school improvement program and its effects on teachers' inquiry teaching behavior.

Summary of the school improvement program and its effects on teachers' inquiry teaching

At each participating school, whole primary school teams (comprising all teachers and their school leaders) received the same training course. Although different operationalizations exist of inquiry-based teaching, the course adhered to the prevailing notion that pupils' inquiry should be guided rather than directed by teachers (see Walan, Mc Ewen, & Gericke, 2016). Thus, participants learned to encourage pupils' own inquiry-oriented ideas, questions, and interests to emerge from their study of school subject matter and to achieve their usual lesson content objectives largely in response to pupils' emerging ideas, questions, and interests (see also Claxton, 2007). To accomplish this, the course comprised two consecutive parts.

First, participants completed the six-months attitude-focused training course by Van Aalderen-Smeets and Walma van der Molen (2015), which aimed at making participants aware of their own attitudes towards inquiry and inquiry teaching, and

to adopt more positive attitudes when necessary. This was deemed important, because teachers' perceptions about such matters are known to guide their classroom practices (e.g., Osborne, Simon & Collins, 2003; Van Aalderen-Smeets & Walma van der Molen, 2015). Second, participants completed a three-months newly developed course, aimed at fostering the skills to develop pupils' attitudes towards epistemic curiosity, implicit ability beliefs, positive effort beliefs, and goal orientations as part of daily lesson objectives. In parallel, participants learned to incorporate more concrete opportunities for pupils' inquiry during their regular lesson activities, such as by stimulating pupils to formulate inquisitive questions about the school subject matter at hand, to make productive connections between different subject domains, and to test their ideas through design experiments. Participants also learned to assess the above lesson objectives by formative pupil feedback.

Post and Walma van der Molen (in press) indicated that the training course significantly improved teachers' inquiry teaching behavior among an intervention group compared to a control group. Immediately after course completion, the teachers had increased their inquiry teaching frequency to an almost daily basis. Also, teachers showed significant improvements in stimulating their pupils' attitudes, beliefs and motivations and integrated more opportunities for pupils to study school subject matter through inquiry. Furthermore, results showed that the teachers sustained their improved practices to one year after course completion (the end of the research project). Please see Table 5.1 for a structured overview of some of the lesson interventions that were implemented by the teachers in the intervention group during the program. For a detailed description of the contents of the course and the study, please see Post and Walma van der Molen (in press).

Table 5.1

Overview of example lesson interventions by the teachers among the intervention group

Component	Example Lesson Interventions
Attitudes towards epistemic curiosity	Role modeling for pupils the use of curiosity as a strategy for exploring novel subject matter, such as by asking follow-up questions or wondering about alternative ideas.
	Inspiring pupils with personal quotes from influential engineers or researchers in the news that underscore the value of curiosity-driven people to society (e.g., Elon Musk).
	Before pupils start to collaborate on a research project, explicitly asking everyone to welcome and encourage each other's curiosity-based ideas and questions about the subject matter at hand.
	After pupils have finished a project, asking them to recall any curiosity-related behavior by their peers that they found inspiring and would like to learn from.
	Rewarding pupils for their inquisitive follow-up questions for investigating lesson content, such as by rewarding them with stickers or displaying their questions on the classroom wall.
Implicit beliefs about the malleability of their ability	Praising pupils' effort, rather than their ability, when they have completed a difficult learning task
	Inspiring pupils to imagine their potential learning ability if they stick with giving it their best effort in school, especially at times when they have failed a test
	Cultivating in pupils a shared notion that making 'mistakes' during learning is an essential and fun part of it, such as by showcasing pupils' most 'promising mistakes' during lesson activities or by displaying these in the classroom.
	Teaching pupils about the plasticity of the human brain and its potential growth through effortful learning, such as by use of online videos about brain development.
	When pupils say that they 'are simply not good at doing maths', helping them to rephrase such perceptions more incrementally, with statements such as 'I am not yet good at maths'.
Positive effort beliefs	Role modeling for pupils the joy and usefulness of putting effort into researching subject matter, such as by joining in on pupils' research assignments.
	Asking pupils to describe their past learning efforts, for example, after they have finished an assignment, in order to discuss resulting learning outcomes.
	Convincing pupils about the causal relation between effort and performance by referring to other kinds of their personal experiences with 'effort', such as in sports.
	Discussing with pupils their perceptions of the value of putting extra effort into learning and asking them to illustrate their answers with personal examples.
	Teaching pupils about the range of effortful actions that may foster their learning performance, such as asking questions or seeking explanations.

Table 5.1 continued

Achievement goal orientations	Encouraging pupils to set learning goals in line with their own personal interests and including these goals on their performance cards.
	Helping pupils not to shy away from difficult learning tasks, but to find a sense of pride in committing to mastering such tasks.
	Setting up friendly competitions that challenge pupils to increase their learning performance, such as who can generate the most inquisitive questions about a particular school subject matter in a limited amount of time.
	Setting aside time in the weekly lesson roster for pupils to explore their own learning interests, either alone or together with peers.
	Inspiring pupils to pick their own academic role models (e.g., engineers, family, celebrities) and to formulate personal goals to become that good themselves.

Present Study

As shown by Post and Walma van der Molen (in press), the school improvement program succeeded in helping teachers create stimulating classroom cultures and practices for pupils' inquiry learning. In the present study, we examined whether these positive changes affected subsequent changes in pupils' (a) attitudes towards epistemic curiosity, (b) implicit ability beliefs, (c) positive effort beliefs and (d) achievement goal orientation motivations. In the following subsections, each of these variables are described in more detail.

Attitude towards epistemic curiosity. Epistemic curiosity is the desire for new intellectual information and is known for driving complex exploratory behavior (Piotrowski, Litman, & Valkenburg, 2014). However, no well-established guidelines or instruments yet exist to promote or assess pupils' epistemic curiosity in school settings (for a review, see Grossnickle, 2016). Post and Walma van der Molen (2018b) were first to propose that, on the basis of attitude theory, pupils' epistemic curiosity can only flourish in classrooms where pupils perceive the epistemic value, pleasure, appreciation, and opportunity of being curiously-minded learners. Together, such cognitive, affective, normative and control perceptions may determine motivations to perform or not perform associated behavior (Ajzen & Fishbein, 1980; Eagly & Chaiken, 1993). Although attitudes are relatively stable psychological constructs, they can be improved through interventions, such as by explicit reflection and discussion activities (Vogel & Wänke, 2016).

Post and Walma van der Molen (2018b) proposed five components of pupils' attitudes towards epistemic curiosity that, on the basis of the Theory of Planned

Behavior (TPB) by Ajzen (2001), may precede pupils' epistemic curiosity in school settings, namely: (1) pupils' perception of the learning value and pleasure of being epistemically curious (Personal Inclination); (2) pupils' perception of the value of epistemically curious people to society (Societal Relevance); (3) pupils' perception of their classmates' negative judgments about them being curious (Fear of Classmates' Negative Judgment); (4) pupils' negative opinions about other people's epistemic curiosity (Negative Opinion); and lastly (5), pupils' perception of their capability to be epistemically curious (Self-Efficacy). For detailed descriptions of these attitude components, including corresponding measurement scales, please see Post and Walma van der Molen (2018b).

Implicit ability beliefs. Another psychological factor that has been suggested to impact pupils' inquiry is pupils' implicit beliefs about the malleability of their 'intelligence'. Dweck (2000) claims that pupils may hold different intelligence, or ability, beliefs. Two opposing beliefs are distinguished, namely the belief that ability is a fixed internal characteristic (Entity Beliefs) and the belief that ability is malleable and can be largely improved through practice and experience (Incremental Beliefs). According to Dweck, and others, these implicit beliefs determine a wide range of achievement-related motivations, such as how pupils approach learning, how they respond to challenges, and how they cope with setbacks (e.g., Burnette, O'Boyle, VanEpss, Pollack, & Finkel, 2013; Mangels et al., 2006). Pupils endorsing an entity belief likely attribute their failure to a lack of innate ability and are thereby prone to disengage from difficult learning tasks or avoid such tasks altogether. Pupils endorsing an incremental belief likely attribute their failure to a lack of practice and are more prone to embrace challenges to grow their competency (Dweck, 2000). Studies show that incremental beliefs can be fostered through lesson interventions, such as by teaching pupils about brain plasticity or by praising their learning effort rather than their ability (e.g., Blackwell et al., 2007; Paunesku, Walton, Romero, Smith, Yeager, & Dweck, 2015).

Positive effort beliefs. Effort beliefs refer to the perceived contingency between one's learning outcomes and one's invested efforts (Dweck & Leggett, 1988). It is well-established in the learning-sciences that effort beliefs play a key role in academic success, including in inquiry-based learning (Claxton & Carr, 2004), as they mediate the link between motivation and performance (Elliot, 1999; Goodman et al., 2011). Pupils' positive effort beliefs toward a learning task must precede their exertion of effort, because the commitment to the task implies the belief that the exerted effort will likely lead to positive outcomes. Thus, research has repeatedly recommended

interventions aimed at teaching pupils about the value of effortful, inquiry-focused thinking as a strategy for attaining improved learning outcomes (e.g., Claxton, 2007).

Achievement goal orientation. Achievement goal orientation represents pupils' purpose for enacting achievement-related behavior (Ames, 1992). Despite discussions about the nature and constellation of achievement goals (e.g., Senko, Hulleman, & Harackiewicz, 2011), research by Elliot and Church (1997) indicates that achievement goals can be understood best by a multiple goals perspective framework, namely: (1) mastery goal orientations, which represent motives to learn for the sake of learning and self-improvement; (2) performance-approach goal orientations, which represent motives to demonstrate competency relative to others; and (3) performance-avoidance goal orientations, which represent motivates to avoid the display of incompetence to others. By this trichotomous model, mastery and performance-approach goals have been associated with a host of positive cognitive, motivational, affective, and behavioral outcomes relevant to inquiry-based learning, such as self-regulated learning (Pintrich, 2003), whereas performance-avoidance goals are associated with no positive outcomes (Fadlelmula, 2010).

Hypotheses

We assessed the above-described variables among the pupils of the participating schools (intervention and control) before (T1), immediately after (T2), and one year after (T3) the school improvement program. As described before, this allowed us to first examine the *influence* of pupils' attitudes towards curiosity and their implicit ability beliefs on their effort beliefs and goal orientation motivations, based on the T1 data. Second, we examined both the short-term and long-term *changes* in pupils' attitude, belief, and motivation scores over time as a result of the school intervention. This longitudinal investigation was believed important, because research suggests that attitude and motivation changes generally require long-term interventions (e.g., Blackwell et al., 2007; Vogel & Wänke, 2016). Since the teachers in the intervention group were still in training during the first year of the program (T1 to T2) and since Post and Walma van der Molen (in press) found that the teachers in the intervention group extended their inquiry teaching practice from T2 to T3, we generally expected to observe the greatest changes in pupils' scores by the end of the program (T3).

Hypothesized relationships among pupils' attitudes, beliefs and motivations

Two theoretical pathways were examined, namely the influence of: (1) attitudes towards epistemic curiosity on effort beliefs and on achievement goal orientation (i.e., on the basis of the TPB by Ajzen, 2001); and (2) implicit ability beliefs on effort beliefs and on achievement goal orientation (i.e., on the basis of the motivation framework described by Blackwell et al., 2007). Please note that, with the separation of these models, we do not imply that attitude towards epistemic curiosity and implicit ability beliefs are necessarily unrelated. Rather, we decided to test two models, because these could be empirically tested more easily, would produce more straightforward interpretations and could lead to more practical recommendations for school intervention programs.

Model 1: The influence of attitudes towards epistemic curiosity on positive effort beliefs and goal orientation motivations. On the basis of the Theory of Planned Behavior by Ajzen (2001), we hypothesized that pupils' attitudes towards epistemic curiosity would directly predict positive effort beliefs and goal orientation motivations. Although the influence of attitude on effort and motivation has been widely accepted (e.g., Ajzen, 2001; Fishbein & Ajzen, 1974), the relation between epistemic curiosity and effort and motivation has been hardly considered (Grossnickle, 2016). Litman (2008) suggested that epistemic curiosity might be positively related to mastery goal orientation, because epistemic curiosity represents the intrinsic desire for intellectual information and thereby evokes a natural focus on mastery goals. In our view, this proposition seems reasonable. However, because the components of pupils' attitudes towards epistemic curiosity stem from different underlying cognitive, affective, normative, and control dimensions of the TPB, we expected that each component would relate differently to effort beliefs and goal orientations.

On the basis of the above literature, we hypothesized that the Personal Inclination component would positively relate to Mastery Goal Orientation and Positive Effort Beliefs, because pupils who perceive the personal value of being curiously-minded would likely seek opportunities for pursuing their own learning interests. We expected the Personal Inclination to be unrelated or negatively related with Performance-Approach Goal Orientation, because pupils who perceive the personal value of curiosity are likely to be less focused on achieving in reference to other people.

We expected the Societal Relevance component to be unrelated or positively related to Mastery Goal Orientation, because the component refers to pupils' perceived value of *other* curiously-minded people to society and not necessarily of their *own*. Thus, we expected this component to positively relate to Performance-Approach Goal Orientation and Positive Effort Beliefs, as these may reflect motives to perform relative to oneself and to others.

We hypothesized the Negative Opinion and Fear of Classmates' Negative Judgement components to be unrelated or negatively related to Mastery Goal Orientation and Performance-Approach Goal Orientation, since pupils with such negative normative perceptions about curiosity-driven thinking likely avoid circumstances that may reveal their inquisitiveness to others. In addition, we expected the Negative Opinion component to be either unrelated or negatively related to Positive Effort Beliefs, because these opinions do not directly relate to pupils' own efforts for learning. In the case of the Fear of Classmates' Negative Judgment component, however, we expected a negative relation with Positive Effort Beliefs, because pupils who hold this fear are likely to avoid showing effortful, curiosity-oriented behavior to their peers. Thus, we expected both attitude components to positively relate to Performance-Avoidance Goal Orientation.

Lastly, we expected the Personal Inclination, Societal Relevance, and Self-Efficacy components to be negatively related to Performance-Avoidance Goal Orientation, because pupils with such positive perceptions intentionally seek opportunities for being inquisitive.

Given the general influence of attitude on effort and motivation, we expected to find the above-described relationships among pupils' attitudes, effort beliefs and goal orientation scores already at the onset (T1) of the program (*Hypothesis 1*).

Model 2: The influence of implicit ability beliefs on positive effort beliefs and goal orientation motivations. We used the motivation framework described by Blackwell et al., (2007) to hypothesize the relationship among implicit ability beliefs and effort beliefs and goal orientations. Such a direct relationship has been hypothesized by others as well, such as Dinger et al., (2013). Hong et al. (1999) showed that pupils who endorse incremental beliefs, are more likely to attribute their learning outcomes to their efforts than pupils who endorse entity beliefs. This link is not surprising, since pupils who endorse incremental beliefs believe in the malleability of their ability through effortful practice (Dweck, 2000). Burnette et al. (2013) indicated that pupils holding incremental ability beliefs also seem predominantly focused on mastery goals, because they view challenging learning

tasks as opportunities to further grow their abilities. Conversely, they focus less on performance-approach and performance-avoidance goals. Pupils who endorse entity beliefs focused more on performance-approach and performance-avoidance goals, because they view their abilities as innately fixed and thus seek tasks to demonstrate their ability or avoid tasks for risk of failure (Burnette e al., 2013).

It should be noted that we are well aware that the association of implicit ability beliefs with motivation is complex and has been described using different mediators (e.g., Tempelaar et al., 2015; Van Aalderen-Smeets & Walma van der Molen, 2018a, 2018b). As discussed before, we ignored these complexities for theoretical and methodological reasons. Given the general influence of implicit ability beliefs on effort and motivation, we hypothesized that at T1, the Incremental Beliefs component positively related to Mastery Goal Orientation and Positive Effort Beliefs, and negatively related to Performance-Avoidance Orientation components. We expected the Incremental Beliefs component to be either unrelated or positive related to Performance-Approach Goal Orientation, because pupils with incremental ability beliefs might still feel driven to demonstrate their achievements to others. Conversely, we expected the Entity Beliefs component to relate positively to Performance-Avoidance Goal Orientation, Performance-Approach Goal Orientation, and Positive Effort Beliefs, but negatively to Mastery Goal Orientation. *Hypothesis 2* comprises the above-described relationships among the variables of Model 2.

Hypothesized effects of the school intervention program on pupil growth

Attitude towards epistemic curiosity. Post and Walma van der Molen (in press) indicated that, after the school improvement program (T2), teachers stimulated pupils to perceive the value, pleasure, appreciation, and opportunities of curiosity-driven thinking. Thus, we expected to observe increases in the Personal Inclination, Societal Relevance, and Self-Efficacy mean scores of pupils in the intervention group and decreases in their Negative Opinion and Fear of Classmates' Negative Judgment mean scores from T1 to T2, while we did not expect such changes in the control group (*Hypothesis 3a*). From T2 to T3, we expected further improved mean scores of the intervention group on these components (*Hypothesis 3b*).

Implicit ability beliefs. In addition, Post and Walma van der Molen (in press) indicated that, after the school improvement program at T2, the teachers in the intervention group encouraged pupils' incremental beliefs throughout their daily teaching, for example by reflective group discussions about the malleability of ability and by praising pupils' learning effort during setback experiences. Given this practice,

we expected increases in the Incremental Belief mean scores and decreases in the Entity Beliefs mean scores of the intervention group from T1 to T2, while we did not expect such changes in the control group (*Hypothesis 4a*). We expected to observe further improved mean scores of the intervention group on these components from T2 to T3 (*Hypothesis 4b*).

Positive effort beliefs. Post and Walma van der Molen (in press) showed that the teachers in the intervention group also learned to stimulate their pupils in articulating positive achievement goals and setting up effortful actions to accomplish these goals. We thus expected increases in the Positive Effort Belief mean scores of the intervention group from T1 to T2, while we did not expect such changes in the control group (*Hypothesis 5a*). We expected to observe further improved mean scores of the intervention group from T2 to T3 (*Hypothesis 5b*).

Achievement goal orientation. From T1 to T2, we also expected increases in the Mastery Goal Orientations and the Performance-Approach Goal Orientation mean scores and decreases in the Performance-Avoidance Goal Orientation mean scores of the intervention group, while we did not expect such changes in the control group (*Hypothesis 6a*). We expected to observe further improved mean scores of the intervention group on these components from T2 to T3 (*Hypothesis 6b*).

Grade level effects. According to Jones, Jones, and Hargrove (2003), high-stake testing and efforts to 'teach to these tests' become increasingly prominent as pupils transition towards the upper grades of primary school. Such teacher-directed and prescriptive practice is believed to offer pupils little encouragement for inquiry and may, consequently, lead them to develop negative attitudes, beliefs, and motivations towards inquiry over time (Post & Walma van der Molen, 2018a, 2018b; Engel, 2015). We thus expected to find relatively less positive scores among the upper-grade pupils on all of our dependent variables at the time of the pretest (T1), compared to the scores of the lower-grade pupils in our sample at T1 (*Hypothesis 7*). In addition, it was believed that the pupils with the most negative mean T1 scores on our dependent variables would show the greatest score improvement over time as a result of the school improvement program, because all of their teachers stimulated positive attitudes, beliefs, and goal orientations to an equal degree (see Post & Walma van der Molen, in press). Thus, we expected to observe increased homogeneity among pupils' attitude, belief and goal orientation mean scores from T1 to T2 and T3. This implied that, given *Hypothesis 7*, we expected significantly greater improvement in the mean scores of the upper-grade pupils on all of our dependent variables during the program

(T1 vs. T2 vs. T3) compared to the mean scores of the lower-grade pupils on these variables (*Hypothesis 8*).

Method

Design and Sample

Six Dutch primary schools volunteered to participate in the study. From these schools, all 4th, 5th and 6th grade pupils (K8, i.e. age 9–12) participated as respondents. Our main reason for focusing on this particular age group was that our survey instruments proved too difficult for younger pupils to comprehend.

A quasi-experimental pretest-posttest control group design with a second posttest was used to assess the effects of the school intervention program on these pupils. The six participating schools were assigned to either one of two groups: (1) the intervention group, for schools that participated in the program; and (2) the control group, for schools that did not participate in the program. Please note that the teachers in the control group participated in the same school intervention program at a later time during the project (school year 2015–2016). However, the examination of the effects of this delayed intervention on pupils' attitudes, beliefs, and motivations fell outside the scope of the current study[2].

During the first research phase (school year 2014–2015), we assessed the effects of the intervention by means of a pretest-posttest control group design (T1 vs. T2), where the changes in pupils' scores in the intervention group were compared to the changes in pupils' scores in the control group. During this first phase, the intervention group consisted of 255 pupils (55.3% female) and the control group consisted of 275 pupils (47.3% female). At T2 (at the end of the same school year), the intervention group remained unchanged, while the control group changed to 273 pupils (47.6% female).

In the second research phase (school year 2015–2016), we examined longitudinal changes in pupils' scores among the intervention group by means of a

---

[2] Post & Walma van der Molen (in press) indicate that the (same) delayed intervention for the control group was incidentally hindered by several school factors (e.g., teachers felt discontent with one of the teacher trainers, new lesson methods were introduced in a school during the program that demanded too much attention of the teachers). As a result, the teachers in the control group generally showed limited improvements in their inquiry teaching practices during the intervention.

one-group longitudinal comparison of results over T1, T2, and T3. For this comparison, we only examined changes in the scores of the cohort of pupils that had remained in the study at all three measurement times (n = 155; 60.6% female). Pupils in Grade 6 that had participated at T1 and T2 were thus not included in this longitudinal investigation, because they had already moved to secondary education at the time of T3.

In addition to the intervention and control schools, we recruited three additional primary schools to serve as a baseline reference group (naive control group) for pupils' scores at T1. Research shows that teachers' interests in inquiry teaching may be positively related to their inquiry teaching behavior, and thereby, pupils' experiences with inquiry learning in school (Lumpe, Czerniak, Haney, & Beltyukova, 2012). Although all six schools that volunteered to participate in our study claimed to have little prior experience with inquiry-based teaching at the onset of the study, we wanted to test this assumption. The naive reference group of schools was recruited from the same area as the six participating schools and comprised 248 pupils (46.0% female).

Measures

We used a survey instrument that comprised four different questionnaires, namely pupils' attitudes towards epistemic curiosity (Post & Walma van der Molen, 2018b), implicit ability beliefs (De Castella & Byrne, 2015), effort beliefs (Blackwell, 2002), and goal orientation motivation (Elliot & McGregor, 2001). In total, the survey instrument consisted of 41 item statements that were consistently measured by means of a four-point Likert-type response option scale. Response options ranged from (1) 'fully disagree', (2) 'somewhat disagree', (3) 'somewhat agree', to (4) 'fully agree' (for a rationale for using a four-point response scale, please see Kulas & Stachowski, 2009). On the basis of recommendations by Schwarz (2008), we made changes to the wording of several item statements of most subscales to improve their comprehensibility for our respondent group. We also randomized the sequencing of the items of each scale.

Data quality assessment. We took several steps to examine the quality of our measurement scales in preparation of our main analyses by means of Mplus version 7.4 (Muthén & Muthén, 1998–2015). We first conducted Exploratory Factor Analysis (EFA) to examine the unrestricted latent factor structure (a maximum number of 10 factors was pre-set). We conducted EFA separately for each scale on the basis of the pretest data (T1) to first establish acceptable item-factor structures, as

each scale represented theoretically distinct psychological constructs. Although our questionnaire was based on pre-validated measures, we had rephrased several item statements, and thus wanted to know whether our revised subscale items would show corresponding latent factor structures. Because our dataset consisted of scores that ranged on a four-point scale, we defined all the items of our dataset as categorical variables and used the robust weighted least squares estimator (WLSMV) and Geomin oblique rotation for all analyses (Muthén & Muthén, 1998–2015).

On the basis of the resulting latent factor structures derived from EFA, for each separate scale, we subsequently conducted Confirmatory Factor Analysis (CFA) with WLSMV to examine how well the pretest data fitted these pre-defined factor structures. To determine model fit, we examined multiple goodness-of-fit indices. We used the absolute model fit estimate Weighted Root Mean Square Residual (WRMR), instead of the conventional χ^2/df fit index, because WRMR is less sensitive to large respondent groups (Prudon, 2015). WRMR should be around 1.0 to indicate acceptable model fit. In addition, we used the Comparative Fit Index (CFI), (3) the Tucker-Lewis Index (TLI), and (4) the RMSEA (Floyd & Widaman, 1995) to examine whether the factor structures of our scales fitted the data. An acceptable fit is indicated by CFI and TLI values between .90 and .95 (good fits > .95), and RMSEA values below .08 (Mueller & Hancock, 2008).

In addition, we followed the guidelines by Carter (2016) to establish the convergent power of each individual factor of a particular scale and its discriminative power relative to all other factors. We calculated Average Variance Extracted (AVE), Average Shared Variance (ASV), and Maximum Shared Variance (MSV) of each individual factor (subscale) based on the coefficients resulting from CFA. The convergent power of a factor is considered satisfactory when AVE is equal to or greater than .50 and the composite reliability (CR) of the factor is greater than .70. Both measures indicate the amount of shared variance among the items that define the factor. Discriminative power is satisfactory when AVE is greater than the Average Shared Variance (ASV) between this factor and all other factors and is greater than the MSV. When results proved acceptable, we computed weighted sum-scores for each factor by averaging respondents' scores on the final set of items that defined the factor. Please see Table 5.1 for an overview of all the items used in the final survey instrument with their corresponding factor loadings as obtained by EFA and CFA on the basis of the pretest data (see Appendix E for Tables A, B, C and D for the complete item statements of the survey instrument). The construct validity indices of each subscale are presented in Table 5.3. As shown in Table 5.3, results indicated that the

convergent and discriminative power of all subscales was acceptable. It should be noted, however, that the above thresholds for model fit and construct validity are simply guidelines and not strict rules (Carter, 2016; Prudon, 2015).

Lastly, we also conducted the above CFA computations and examinations separately on our T2 data of both groups and T3 data on the remaining cohort of pupils among the intervention group. As described, one could expect changes in the relationships among our dependent variables as a result of pupils' attitude, belief, and motivation development. We thus tested whether the pre-defined latent factor structures found in the pretest data (T1) remained consistent across the two years of the program (T2 and T3), in order to allow reliable comparative examinations of pupils' mean scale scores over time. Results showed that all of our pre-defined factor structures at T1 remained consistent at T2 and T3, as indicated by acceptable model fit indices. Please see Table 5.4 for the model fit indices of the different scales at all three measurement times.

Attitude towards epistemic curiosity. Attitudes towards epistemic curiosity was assessed by means of the Attitudes Towards Epistemic Curiosity scale taken from the Children's Images and Attitudes of Curiosity (CIAC) questionnaire by Post and Walma van der Molen (2018b). The scale consisted of 17 items in total; 4 items for the Personal Inclination subscale, 3 items for the Societal Relevance subscale, 3 items for the Negative Opinion subscale, 3 items for the Fear of Classmates' Negative Judgment subscale, and 4 items for the Self-Efficacy subscale. As shown in Table 5.2 and 5.4, results indicated good factor loadings and model fits during the entire program.

Implicit ability beliefs. Our implicit ability beliefs scale was based on the scale by De Castella and Byrne (2015), which comprised 4 items for the Incremental Beliefs subscale and 4 items for the Entity Beliefs subscale. De Castella and Byrne (2015) revised the items of the original implicit intelligence beliefs scale by Dweck (2000) to measure respondents' beliefs about the malleability of their *own* intelligence, instead of the malleability of intelligence in *general*. In addition, we replaced the term 'intelligence' with the term 'ability' in the item statements of the revised implicit intelligence scale by De Castella and Byrne (2015). The term 'intelligence' by Dweck (2000) has been critiqued for being too ambiguous and too complex (Van Aalderen-Smeets, Walma van der Molen, & Xenidou-Dervou, 2018b). We agree that these issues could result into varying interpretations by the primary school pupils in our sample when rating their ability beliefs (e.g., Does the term

'intelligence' also imply 'factual knowledge'?). To prevent these potential psychometric issues, in the current study, we thus used the word 'ability' instead.

In line with De Castella and Byrne (2015), results of the EFA and the CFA indicated a two-factor solution at T1, T2 and T3, showing good item factor loadings, that corresponded with our hypothesized factor structure (see Table 5.2 and 5.4).

Effort beliefs. The Positive Effort Beliefs scale used in the present study was taken from the Effort Belief scale by Blackwell (2002) and comprised 4 positive items (see also Blackwell et al., 2007). As expected, EFA results indicated that all four items loaded on a single factor. CFA confirmed that this factorial structure derived from the EFA fitted the data, as indicated by good model fit indicates at T1, T2 and T3. Please see Table 5.2 and 5.4 for these results.

Achievement goal orientation. Lastly, Goal Orientation Motivations was operationalized by the AGQ items from Elliot and McGregor (2001), which distinguishes 5 items for the Mastery Goal Orientation subscale, 5 items for the Performance-Approach Goal Orientation subscale, and 5 items for the Performance-Avoidance Goal Orientation subscale. However, EFA results revealed a two-factor solution, indicating that the five items belonging to the Mastery Goal Orientation subscale and the five items of the Performance-Avoidance Goal Orientation subscale all loaded on their respective factors. The five items belonging to the Performance-Approach Goal Orientation subscale, however, showed poor loadings and cross-loadings. We thus removed these items from our data set, leaving 10 items that were inputted into a CFA with the pre-defined factor structures as derived from the EFA. CFA confirmed that this two-factorial structure fitted the data, as indicated by good model fit indices at T1, T2 and T3 (see Table 5.2 and 5.4).

Table 5.2
Items used in the survey instrument with corresponding factor loadings as obtained by exploratory and confirmatory factor analysis

Item code	EFA[a]										CFA[b]									
	PI	SR	NO	FJ	SE	IB	EB	PE	MA	PA	PI	SR	NO	FJ	SE	IB	EB	PE	MA	PA
PI1	.67										.63									
PI2	.34										.75									
PI3	.50										.72									
PI4	.69										.75									
SR1		.72										.70								
SR2		.70										.71								
SR3		.72										.78								
NO1			.78										.59							
NO2			.59										.76							
NO3			.60										.71							
FJ1				.77										.71						
FJ2				.72										.78						
FJ3				.71										.73						
SE1					.55										.74					
SE2					.57										.73					
SE3					.75										.71					
SE4					.61										.77					
IB1						.53										.58				

Table 5.2 continued

IB2	.68	.75
IB3	.66	.71
IB4	.79	.83
IB5	.69	.75
EB1	.72	.63
EB2	.72	.76
EB3	.66	.70
EB4	.77	.81
EB5	.69	.83
PE1	.83	.83
PE2	.78	.78
PE2	.67	.67
PE4	.63	.63
MA1	.72	.83
MA2	.71	.73
MA3	.69	.77
MA4	.71	.78
MA5	.73	.77
PA1	.70	.83

Table 5.2 continued

PA2	.60	.54
PA3	.70	.85
PA4	.75	.73
PA5	.80	.79

<u>Notes:</u>

[a] Exploratory factor analysis (EFA) conducted with weighted least squares estimation with means and variance adjusted and Geomin rotation on the T1 pretest data (*n* = 530). Values represent rotated factor loadings.

[b] Confirmatory factor analysis (CFA) conducted with weighted least squares estimation with means and variance adjusted on the pretest data (T1). Values represent factor loadings.

Factor loadings are only displayed for items with loadings > .30 on their expected factors and with cross loadings > .15 difference from the item's greatest factor loading.

PI = Personal Inclination, SR = Societal Relevance, NO = Negative Opinion, FJ = Fear of Classmates' Negative Judgment, SE = Self-Efficacy, IB = Incremental Beliefs, EB = Entity Beliefs, PE = Positive Effort Beliefs, MA = Mastery Goal Orientation, PA = Performance Avoidance Goal Orientation. The complete item statements of the survey instrument are listed in "Appendix E".

Table 5.3

Factor correlations, CR, AVE, MSV, and ASV for the subscales as obtained by confirmatory factor analysis on the basis of the T1 pretest data (n = 530)

	Attitudes Towards Epistemic Curiosity					Implicit Ability Beliefs		Effort Beliefs	Goal Orientation Motivations	
	PI	SR	NO	FJ	SE	IB	EB	PE	MA	PA
Attitudes Towards Epistemic Curiosity										
Personal Inclination (PI)										
Societal Relevance (SR)	.75*									
Negative Opinion (NO)	−.38*	−.33*								
Fear of Classmates' Negative Judgment (FJ)	−.03	.14*	.56*							
Self-Efficacy (SE)	.83*	.59*	−.33*	−.07						
Implicit Intelligence Beliefs										
Incremental Beliefs (IB)	.54*	.49*	−.29*	−.09	.43*					
Entity Beliefs (EB)	−.23*	−.24*	.40*	.23*	−.09	−.71*				
Effort Beliefs										
Positive Effort Beliefs (PE)	.71*	.55*	−.32*	−.19*	.61*	.57*	−.28*			
Goal Orientation Motivations										
Mastery Goal Orientation (MA)	.72*	.51*	−.32*	−.19*	.61*	.57*	−.28*			
Performance-Avoidance Goal Orientation (PA)	−.40*	−.39*	.53*	.35*	−.28*	−.25*	−.38*	−.53*	−.55*	

Table 5.3 continued

Composite Reliability (CR)	.80	.78	.73	.79	.83	.85	.87	.82	.88	.87
Average Variance Extracted (AVE)	.63	.67	.60	.68	.67	.66	.69	.67	.73	.70
Maximum Shared Variance (MSV)	.69	.57	.31	.31	.69	.50	.50	.64	.64	.30
Average Shared Variance (ASV)	.32	.23	.17	.07	.24	.22	.12	.29	.31	.17

*Factor correlation is statistically significant at $p < .05$

Table 5.4

Model fits of the different scales as obtained by CFA at T1, T2 and T3.

	WRMR	CFI	TLI	RMSEA
Attitudes Towards Epistemic Curiosity				
T1	1.0	.97	.96	.05
T2	1.1	.96	.95	.05
T3	.93	.96	.95	.07
Implicit Intelligence Beliefs				
T1	.78	.99	.98	.06
T2	1.0	.97	.96	.08
T3	.99	.94	.92	.14
Positive Effort Beliefs				
T1	.52	.99	.98	.10
T2	.20	1.0	1.0	.00
T3	.45	.99	.97	.16
Goal Orientation Motivations				
T1	.97	.98	.98	.07
T2	1.0	.98	.98	.08
T3	.75	.99	.98	.08

T1 (intervention and control group): $n = 530$, T2 (intervention and control group): $n = 528$, T3 (remaining cohort of the intervention group): $n = 155$.

Procedure

A paper-and-pencil version of the complete questionnaire was administered by a research assistant to all the pupils at T1 (October 2014), T2 (June, 2015), and T3 (June, 2016), in their own classrooms during normal school hours. First, the assistant introduced the data collection procedure to the whole class. The pupils were given the time needed to complete the questionnaire (about 20 minutes). If a child did not understand a particular item, the assistant provided feedback individually. The principal author of this study explained the objectives, time requirements, and nature of the data collection procedure to the school management, teachers, and parents several weeks prior the start of the data collection. Informed consent was obtained from the parents of all the pupils who participated in the study in accordance with the ethical guidelines of our university.

Data analyses

Investigating the relations among attitude, beliefs and motivation. We used a Structural Equation Modeling (SEM) approach in Mplus to investigate the hypothesized structural relations among attitudes, beliefs, and motivations, as specified before in *Hypotheses 1* and *2*. We thus tested two structural models using SEM analysis, that describe these hypothesized relationships among our variables (i.e., latent factors). Therefore, we first tested both hypothesized models on the basis of the T1 pretest data of the intervention group and the control group ($n = 530$).

Investigating pupils' attitude, belief and motivation development over time. The effects of the intervention on pupils' development were mostly analyzed by means of two-way, GLM repeated measures MANOVA's with Bonferroni correction using SPSS version 24. In doing so, pupils' attitude, belief and motivation scores were used as the dependent, within-group variables, while condition and grade level were used as the between-group variables. To gain insight into the effects of the intervention on each separate subscale, we performed post-hoc univariate analyses within each experimental group separately (intervention and control group) using paired *t*-tests.

Preliminary data checks

Preliminary data checks revealed that the percentage of missing data in our total sample was acceptable (2.24%). In addition, one-way MANOVA test results indicated statistically significant differences between the mean pretest scores of the intervention group and the control group on several of our dependent variables at T1, $\Lambda = .98$, $F(5, 488) = 2.20$, $p = .05$, $\eta^2 = .02$. However, these mean score differences were marginal. Also, results indicated small, but statistically significant differences between the mean pretest scores of one or both of the experimental groups and the baseline reference group on most of our variables ($p < .05$). Overall, the pupils in the experimental groups held somewhat more positive attitude, belief and motivation scores at the onset of the study compared to the reference group. Please see Table 5.5 for the mean pretest scores of the intervention group, the control group, and the baseline reference group on each subscale.

Results

Relationships among pupils' attitudes, beliefs and motivations

We first tested our hypothesized relationships among pupils' attitudes, beliefs, and motivation scores by Model 1 and 2 on the basis of the T1 pretest data of the intervention group and the control group (n = 530).

  Model 1: The influence of attitudes towards epistemic curiosity on positive effort beliefs and goal orientation motivations. At T1, the observed statistical relationships of the Attitude towards Epistemic Curiosity factors on the Positive Effort Beliefs (R^2 = .54, p < .01), Mastery Goal Orientation (R^2 = .58, p < .01), and Performance-Avoidance Goal Orientation factors (R^2 = .37, p < .01) largely confirmed our predictions as specified in Model 1 (*Hypothesis 1*, largely supported), WRMR = 1.18, CFI = .96, TLI = .95, RMSEA = .05. Please see Figure 5.1a for the regression pathway coefficients among the variables of Model 1. Notably, in contrast to our predictions, no statistically significant relation was found between the Societal Relevance and Positive Effort Beliefs factors nor did we find any statistically significant relation between the Self-Efficacy factor and the Positive Effort Beliefs, Mastery Goal Orientation, and Performance-Avoidance Goal Orientation factors.

  Model 2: The influence of implicit ability beliefs on positive effort beliefs and goal orientation motivations. The observed statistical relationships between the Implicit Ability Beliefs factors and the Positive Effort Beliefs (R^2 = .35, p < .01), the Mastery Goal Orientation (R^2 = .38, p < .01) and the Performance-Avoidance Goal Orientation factors (R^2 = .14, p < .01) also largely confirmed our predictions as specified in Model 2 (*Hypothesis 2*, largely supported), WRMR = 1.26, CFI = .95, TLI = .95, RMSEA = .06. Please see Figure 5.1b for the regression pathway coefficients among the variables of Model 2.
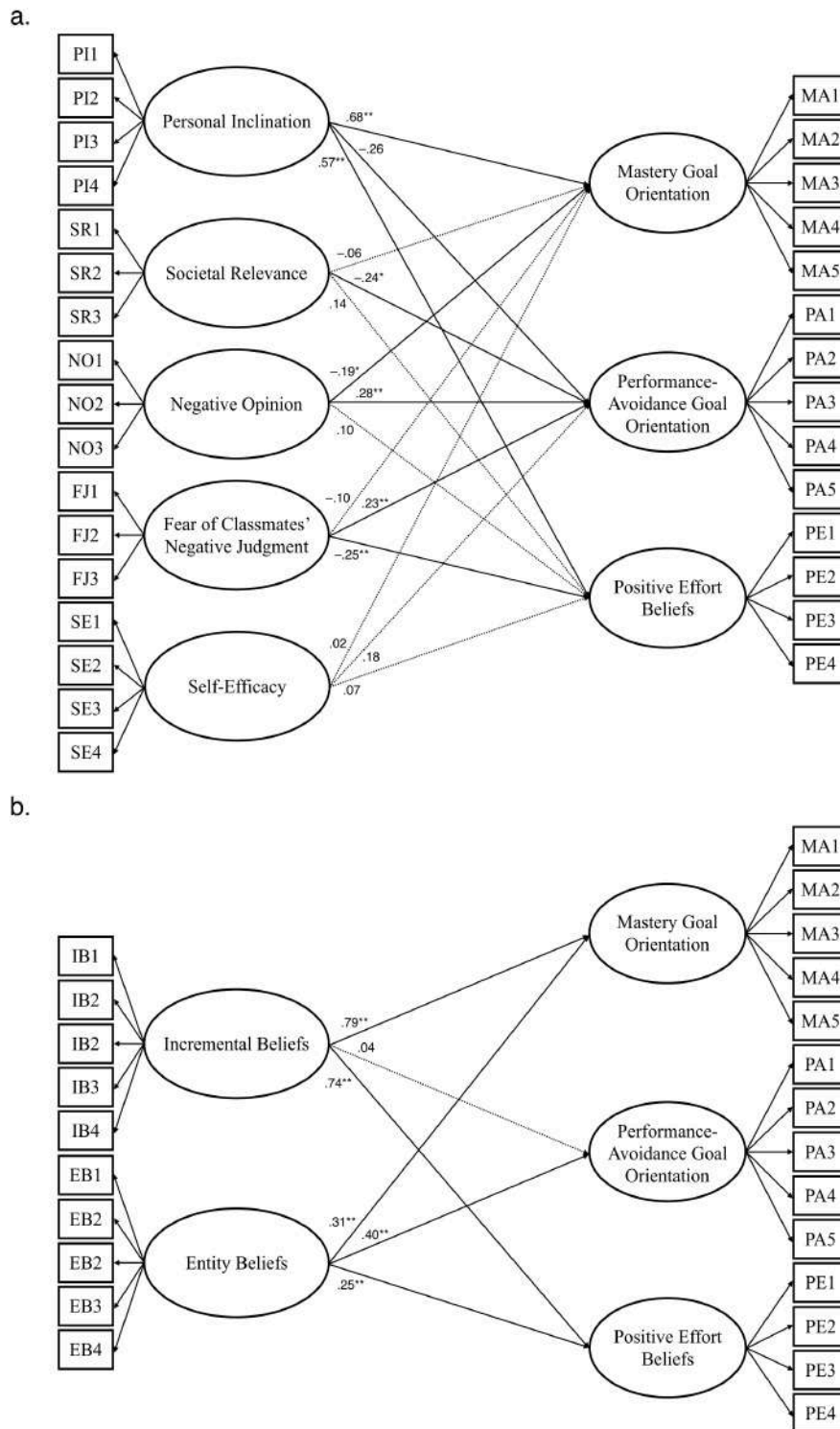
Figure 5.1. Structural mediation models of the influence of (a) Attitudes Towards Epistemic Curiosity and (b) Implicit Ability Beliefs on Positive Effort Beliefs and Achievement Goal Orientations on the T1 pretest data (*n* = 530). Reported β's are the standardized estimates. Circles present latent variables and rectangles represent measured variables. Dotted lines represent non-significant relationships.

* *p* ≤ .05, ** *p* ≤ .01

Effects of the school intervention program on pupil growth

Next, we examined the general effects of the school intervention program on pupils' attitudes, beliefs, and motivations separately for the first research phase (T1 vs. T2) and the second research phase (T1 vs. T2 vs. T3). Because of space limitations of this paper, we present our post-hoc *t*-test results in abbreviated form in table format and only describe statistically significant test results. Please see Table 5.5 for all our repeated measures test results, including the mean pretest and posttest scores of the intervention group and the control group on each subscale per measurement time (T1, T2, and T3). Please see Figure 5.2 for a graphical representation of (the changes in) pupils' mean scores over time on all of our dependent variables for the intervention and control group.

Results of the first phase of the program

Attitude towards epistemic curiosity. MANOVA results, using Wilks' statistic, indicated significant differences in pupils' mean pretest scores between the 4th, 5th and 6th grade level groups across the intervention group and the control group (n = 530), Λ = .93, $F(10, 1448)$ = 5.31, p < .01, η^2 = .04. Post-hoc univariate analyses indicated that these differences only existed for the Personal Inclination ($F(2, 280)$ = 11.19, p < .01, η^2 = .03), Societal Relevance ($F(2, 288)$ = 15.35, p < .01, η^2 = .04), and Self-Efficacy ($F(2, 282)$ = 14.61, p < .01, η^2 = .04) components. As can be seen in Table 5.5, these mean pretest scores were significantly higher for the pupils in the lower grades compared to the pupils in the higher grades. These results thus largely supported *Hypothesis 7*.

Results of a repeated measures MANOVA only indicated a significant higher-order interaction effect between time, condition and grade level on the five attitude components, Λ = .90, $F(10, 904)$ = 4.80, p < .01, η^2 = .05. As shown in Table 5.5, the mean scores of the 5th and 6th grade pupils significantly increased on most of the attitude components from T1 to T2, while the mean scores of the lower-grade pupils on these components remained statistically unchanged. In line with our predictions (*Hypothesis 3a* and *8*), this result thus indicates that the program succeeded in improving the attitudes of particularly the higher-grade pupils in our sample.

After the intervention, at T2, MANOVA results still indicated significant mean score differences among the three grade level groups on most attitude components, Λ = .83, $F(10, 476)$ = 4.59, p < .01, η^2 = .09. However, as can be seen in Table 5.5, these mean posttest scores did no longer show the same general decline from the lower to the higher grades, as observed at T1.

Implicit ability beliefs. MANOVA results indicated that at T1, across both conditions, pupils' Incremental Beliefs ($F(2, 160) = 7.30$, $p < .01$, $\eta^2 = .03$) mean scores were significantly higher for the pupils in the lower grades compared to the pupils in the higher grades, $\Lambda = .94$, $F(10, 964) = 7.25$, $p < .01$, $\eta^2 = .03$ (see Table 5.5). However, this score pattern was not observed for the Entity Beliefs component across both conditions, $F(2, 160) = 1.10$, $p = .33$, $\eta^2 = .01$. These results thus partially support *Hypothesis 7*.

As shown in Table 5.5, repeated measures MANOVA results indicated that pupils' Incremental Beliefs mean scores in the intervention group had significantly increased and their Entity Belief mean scores had decreased from T1 to T2, while the mean scores of the control group on these components had remained statistically unchanged ($p < .05$), $\Lambda = .98$, $F(2, 462) = 4.39$, $p = .01$, $\eta^2 = .02$. No higher-order interaction effect was found between time, condition, grade level and the two implicit ability belief components, $\Lambda = .99$, $F(4, 924) = .77$, $p = .54$, $\eta^2 = .00$. This result thus supported *Hypothesis 4a*, but rejected *Hypothesis 8* in the case of pupils' implicit ability beliefs.

MANOVA results on the posttest data (T2) of the intervention group indicated no significant differences between the mean scores of the 4th, 5th, and 6th grade pupils, $\Lambda = .96$, $F(4, 494) = 2.26$, $p = .06$, $\eta^2 = .02$. These results thus indicate that the school intervention program improved pupils' implicit ability beliefs, from T1 to T2, across all three grade levels.

Positive effort beliefs. ANOVA results indicated statistically significant differences between the mean pretest scores of the 4th, 5th and 6th grade pupils across both conditions, $F(2, 484) = 10.66$, $p < .01$, $\eta^2 = .04$. As shown in Table 5.5, the mean pretest scores of the 5th and 6th grade pupils in both conditions were significantly lower than for the 4th grade pupils. Thus, once again, our mean pretest score difference analysis provided evidence to support *Hypothesis 7*.

Repeated measures ANOVA results revealed a statistically significant, but small interaction effect between time, condition, grade level, $F(2, 463) = 3.76$, $p < .01$, $\eta^2 = .02$. Table 5.5 shows that pupils' Positive Effort Beliefs mean scores of the 5th and the 6th graders in the intervention group had increased from T1 to T2, while the mean scores of the 4th, 5th, and 6th grade pupils in the control group had remained statistically unchanged. Thus, during the first research phase, the school intervention program succeeded in improving the positive effort beliefs of only the 5th and 6th grade pupils. This result thus supported our predictions (*Hypothesis 5a* and *8*).

After the intervention, at T2, a one-way ANOVA conducted on the posttest data (T2) of the intervention group no longer indicated statistically significant differences between the mean scores among the three grade level groups, $F(2, 248) = 2.24$, $p = .11$, $\eta^2 = .02$.

Achievement goal orientation. MANOVA results indicated statistically significant differences between the mean pretest scores of the 4th, 5th, and 6th grade pupils across both conditions, $\Lambda = .94$, $F(4, 936) = 7.99$, $p < .01$, $\eta^2 = .03$. Post-hoc univariate tests indicated that these differences existed for both the Mastery Goal Orientation ($F(2, 469) = 15.12$, $p < .01$, $\eta^2 = .06$) and the Performance-Avoidance Goal Orientation ($F(2, 469) = 3.69$, $p = .03$, $\eta^2 = .02$) components. Overall, the goal orientation mean scores of the pupils in both groups were relatively increased for the lower-grade pupils compared to the higher-grade pupils, which thereby provided further support for *Hypothesis 7*.

Repeated measures MANOVA results indicated a significant, but small interaction effect between time and condition for the achievement goal orientation scale, $\Lambda = .97$, $F(2, 445) = 7.69$, $p < .01$, $\eta^2 = .03$. As shown in Table 5.5, pupils' Mastery Goal Orientation mean score of the intervention group had significantly increased and their Performance-Avoidance Goal Orientation mean score had significantly decreased from T1 to T2, while the mean scores of the control group on these components had remained statistically unchanged. Furthermore, the MANOVA results indicated a significant higher-order interaction effect between time, condition and grade level, $\Lambda = .93$, $F(4, 890) = 7.73$, $p < .01$, $\eta^2 = .03$. Table 5.5 shows that the school intervention group improved pupils' goal orientation motivations during the first research phase, especially among the 5th and 6th grade pupils. This result thus provided support for *Hypotheses 6a* and *8*.

After the intervention, at T2, MANOVA results still indicated statistically significant differences between the mean posttest scores of the 4th, 5th, and 6th grade pupils in the intervention group, $\Lambda = .88$, $F(4, 430) = 7.02$, $p < .01$, $\eta^2 = .06$. However, as can be seen in Table 5.5, pupils' mean scores did not show the same general decline in scores from the lower to the higher grades anymore, as observed at the time of the pretest.

Table 5.5

Mean scores and standard deviations per group on each subscale at T1, T2 and T3 and interactions with condition and grade level

	Intervention group						Control group				Reference		T1 vs T2				T2 vs T3				T1 vs T3			
	T1		T2		T3		T1		T2		T1		Time x Condition		Time x Condition x Grade		Time		Time x Grade		Time		Time x Grade	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	p	η²	p	η²	p	η²	p	η²	p	η²	p	η²
Attitude Towards Epistemic Curiosity																								
Personal Inclination																								
Grade 4	3.02	.65	2.82	.56	–	–	2.72	.64	3.05[a]	.64	2.70	.65	.85	.00	**.00**	**.05**	.19	.01	.27	.01	**.02**	**.04**	**.00**	**.06**
Grade 5	2.70	.64	3.06[a]	.57	2.97	.56	2.78	.68	2.70	.58	2.63	.49												
Grade 6	2.49	.69	2.78[a]	.53	3.05[c]	.58	2.69	.65	2.80	.60	2.46	.56												
Total	2.71	.69	2.88[a]	.56	3.02[c]	.57	2.73	.65	2.85[a]	.62	2.58	.56												
Societal Relevance																								
Grade 4	2.95	.60	2.81	.69	–	–	2.89	.67	2.72	.74	2.65	.61	.27	.00	.92	.00	.37	.01	.38	.01	.97	.00	**.03**	**.03**
Grade 5	2.73	.59	2.78	.58	2.85	.54	2.80	.72	2.69	.64	2.72	.52												
Grade 6	2.42	.72	2.65[a]	.56	2.87	.57	2.66	.61	2.70	.65	2.50	.61												
Total	2.66	.68	2.74	.61	2.86	.55	2.78	.67	2.70	.68	2.62	.59												
Negative Opinion																								
Grade 4	2.01	.74	2.13	.71	–	–	2.00	.64	2.00	.77	1.87	.73	.06	.01	**.01**	**.02**	**.01**	**.04**	**.02**	**.04**	**.00**	**.08**	.96	.00
Grade 5	1.92	.67	1.62[a]	.51	1.76[bc]	.59	1.75	.69	1.92	.68	1.87	.58												
Grade 6	2.00	.74	1.82[a]	.63	1.62[c]	.66	1.93	.65	1.84	.67	1.90	.66												
Total	1.97	.72	1.84[a]	.65	1.68[bc]	.63	1.90	.68	1.92	.71	1.88	.65												
Fear of Classmates' Negative Judgment																								
Grade 4	1.92	.95	2.08	.65	–	–	1.85	.71	1.93	.79	1.70	.60	.07	.01	**.01**	**.02**	**.00**	**.10**	**.03**	**.03**	**.00**	**.10**	.66	.00
Grade 5	1.94	.74	1.64[a]	.58	1.60[bc]	.57	1.68	.62	1.88[a]	.66	1.92	.59												
Grade 6	1.90	.70	1.92	.69	1.52[c]	.67	1.76	.66	1.83	.60	1.77	.65												
Total	1.92	.77	1.87	.67	1.56[bc]	.63	1.76	.66	1.88[a]	.69	1.80	.62												

Table 5.5 continued

	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Self-Efficacy												
Grade 4	2.82	.62	2.58[a]	.58	–	–	2.46	.69	2.72[a]	.71	2.50	.68
Grade 5	2.47	.63	2.79[a]	.57	2.69	.61	2.60	.65	2.42	.62	2.24	.52
Grade 6	2.34	.64	2.53	.53	2.81[c]	.73	2.39	.68	2.56	.63	2.11	.49
Total	2.51	.66	2.63	.57	2.76	.67	2.48	.67	2.57	.66	2.27	.58
p / η^2	.94	.00	.00	.04	.36	.01	.47	.00	.14	.02	.00	.06
Implicit Ability Beliefs												
Incremental Beliefs												
Grade 4	3.03	.68	3.10	.52	–	–	3.14	.53	3.19	.56	3.02	.46
Grade 5	2.98	.50	3.23[a]	.49	3.11	.50	3.08	.46	3.01	.51	3.13	.36
Grade 6	2.82	.65	3.06[a]	.50	3.18[c]	.49	2.90	.61	2.96	.57	3.04	.42
Total	2.93	.62	3.13[a]	.50	3.15[c]	.50	3.03	.54	3.05	.55	3.06	.42
p / η^2	.03	.01	.27	.01	.83	.00	.61	.00	.02	.04	.23	.01
Entity Beliefs												
Grade 4	2.20	.65	1.98[a]	.65	–	–	2.11	.62	2.00	.55	2.10	.55
Grade 5	2.18	.61	2.03	.56	1.98[c]	.59	1.92	.56	2.03	.56	2.00	.50
Grade 6	2.15	.72	2.03	.66	1.87[bc]	.57	2.00	.62	2.13	.68	1.88	.50
Total	2.18	.66	2.01[a]	.62	1.92[c]	.58	2.01	.61	2.06	.60	1.98	.52
p / η^2	.00	.02	.61	.00	.20	.01	.24	.01	.00	.10	.65	.00
Effort Beliefs												
Positive Effort Beliefs												
Grade 4	3.14	.60	2.99	.64	–	–	3.08	.69	3.17	.54	3.11	.53
Grade 5	2.84	.50	3.16[a]	.54	3.07	.57	3.03	.54	2.96	.56	3.00	.42
Grade 6	2.81	.58	3.01[a]	.53	3.03[bc]	.52	2.80	.62	2.86	.60	2.82	.40
Total	2.91	.57	3.06[a]	.57	3.05	.54	2.96	.63	2.99	.58	2.96	.46
p / η^2	.31	.00	.00	.02	.53	.00	.07	.02	.51	.00	.06	.03
Achievement Goal Orientations												
Mastery Goal Orientation												
Grade 4	3.28	.55	3.13	.53	–	–	3.19	.53	3.27	.54	3.13	.62
Grade 5	2.92	.52	3.39[a]	.48	3.16	.55	3.15	.57	2.99	.54	3.01	.41
Grade 6	2.82	.63	3.12[a]	.47	3.18[bc]	.53	2.93	.62	2.94	.58	2.98	.44
Total	2.98	.60	3.27[a]	.54	3.17	.54	3.08	.59	3.06	.57	3.03	.49
p / η^2	.00	.02	.00	.05	.18	.01	.06	.02	.22	.01	.01	.05
Performance-Avoidance Goal Orientation												
Grade 4	1.98	.70	2.21	.68	–	–	1.92	.64	1.99	.68	1.98	.74
Grade 5	2.38	.74	1.71[a]	.58	1.95[b]	.64	1.89	.55	2.07	.62	1.98	.53
Grade 6	2.25	.66	1.99[a]	.67	1.86[c]	.65	2.00	.61	1.96	.64	2.13	.55
Total	2.22	.71	1.96[a]	.67	1.90[c]	.64	1.94	.60	2.00	.65	2.04	.61
p / η^2	.00	.02	.00	.04	.45	.00	.00	.06	.00	.08	.01	.05

167

Notes for Table 5.5:

Mean scores could range between 1 (totally disagree) and 4 (totally agree).

[a]Significant difference (*p* < .05) between T1 and T2 analyzed with a paired *t*-test.

[b]Significant difference (*p* < .05) between T2 and T3 analyzed for the applicable pupil cohort with a paired *t*-test (*n* = 155).

[c]Significant difference (*p* < .05) between T1 and T3 analyzed for the applicable pupil cohort with a paired *t*-test (*n* = 155).

p-values printed in bold indicate a significant (interaction) effect
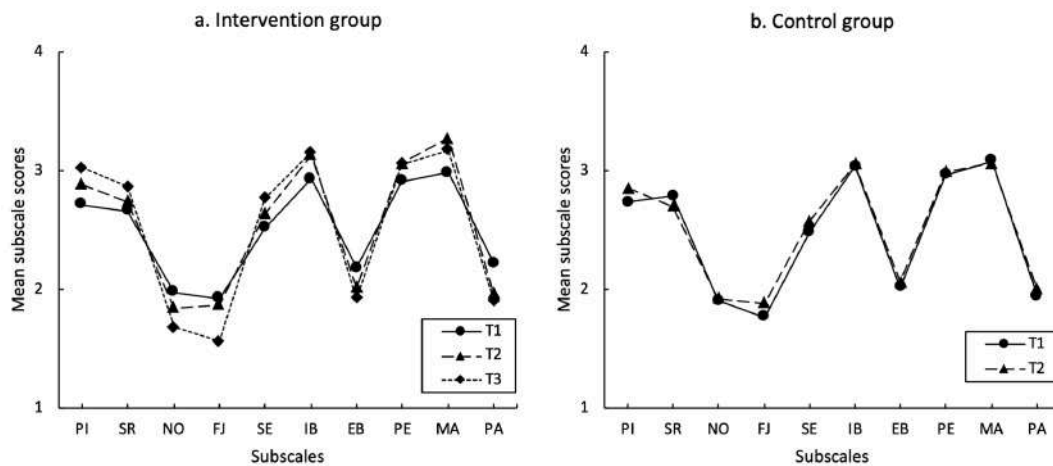
Figure 5.2. Mean subscales scores before (T1), during (T2), and at the end (T3) of the school intervention program on each subscale: Personal Inclination (PI), Societal Relevance (SR), Negative Opinion (NO), Fear of Classmates' Negative Judgment (FJ), Self-Efficacy (SE), Incremental Beliefs (IB), Entity Beliefs (EB), Positive Effort Beliefs (PE), Mastery Goal Orientation (MA), and Performance-Avoidance Goal Orientation (PA). The figure shows the scores of (a) the intervention group (T1 and T2: *n* = 255; T3: *n* = 155) and (b) the control group (T1 and T2: *n* = 275). The connecting lines between the dots are only for presentation purposes and do not depict actual data.

Results of the second phase of the program

Lastly, we examined the degree to which the intervention group who had completed the survey at T1, T2, and T3 (*n* = 155) further improved their attitude, belief, and motivation scores to the end of the research project (T3). We examined pupil growth in two consecutive steps, namely first from T2 to T3 and second from T1 to T3.

Attitude towards epistemic curiosity. Repeated measures MANOVA, using grade level (note: a newly created cohort variable) as the between-subjects factor and

time as a within-subjects factor, indicated a large main effect of time across all five attitude components from T2 to T3, $\Lambda = .89$, $F(5, 141) = 3.39$, $p = .01$, $\eta^2 = .11$. Table 5.5 shows that pupils' Negative Opinion and Fear of Classmates' Negative Judgment mean scores in the 5th grade of the intervention group had significantly decreased from T2 to T3. Thus, during the second phase, the school intervention program only succeeded in further improving pupils' perceptions about the social classroom norms of curiosity-oriented learning. This result thus partially supported *Hypothesis 3b*.

Results revealed no higher-order interaction effect between time and grade level, $\Lambda = .93$, $F(5, 141) = 2.02$, $p = .08$, $\eta^2 = .07$. Yet, as presented in Table 5.5, post-hoc univariate tests did indicate statistically significant differences in pupils' attitude development among the grade level groups from T2 to T3, as the 5th grade pupils at T2 appeared to show greater decreases in their mean scores from T2 to T3, compared to the mean score differences of the 4th graders from T2 to T3, for the Negative Opinion and Fear of Classmates' Negative Judgment components. MANOVA results indicated no significant mean score differences between the 5th and 6th grade level groups across the five attitude components at T3, $\Lambda = .98$, $F(5, 145) = .67$, $p = .65$, $\eta^2 = .02$.

In addition, repeated measures MANOVA results showed that, across all three grade level groups, the pupils in the intervention group indicated increased attitude scores on the final posttest (T3) compared to their pretest scores (T1), $\Lambda = .86$, $F(5, 133) = 4.21$, $p < .01$, $\eta^2 = .14$. However, as shown in Table 5.5, univariate tests showed that this effect only occurred in the case of the Personal Inclination, Negative Opinion, and Fear of Classmates' Negative components.

The MANOVA results also indicated a modest, significant interaction effect between time and grade level across all five attitude components from T1 to T3, $\Lambda = .91$, $F(5, 133) = 2.74$, $p = .02$, $\eta^2 = .09$. Table 5.5 shows that, in particular, the 5th grade pupils significantly increased in their Personal Inclination, Societal Relevance, and Self-Efficacy mean scores from T1 to T3, while the mean scores of the 4th grade pupils did not change. Altogether, these results thus indicate that, across the two years of the program, the program proved most beneficial for the higher-grade pupils in fostering positive attitudes towards epistemic curiosity (supporting *Hypothesis 8*).

Implicit ability beliefs. Repeated measures MANOVA results indicated no significant main effect of time on the two implicit ability beliefs components during the second phase of the program (T2 vs. T3), $\Lambda = .98$, $F(2, 148) = 1.24$, $p = .29$, $\eta^2 = .02$, neither did we find an interaction effect between time and grade level, $\Lambda = .98$, $F(2, 148) = 1.43$, $p = .24$, $\eta^2 = .02$. At T3, no significant mean score differences existed

between the 5th and 6th grade pupils on either implicit ability belief component, as indicated by a MANOVA across both grade level groups, Λ = .99, $F(2, 151)$ = .71, p = .49, η^2 = .01. Thus, as opposed to *Hypotheses 4b* and *8*, the intervention group did not further improve its implicit ability beliefs during the second phase of the program, but did sustain their improved beliefs from the first phase to the end of the second phase of the program (T3).

However, repeated measures MANOVA results did indicate a significant main effect of time across the two implicit ability beliefs components from T1 to T3, Λ = .90, $F(2, 136)$ = 7.55, p < .01, η^2 = .10. Table 5.5 shows that, irrespective of grade level, pupils' Incremental Beliefs mean scores had increased and their Entity Beliefs mean scores had decreased during this time. No higher-order interaction effect between time and grade level was found, Λ = .99, $F(2, 136)$ = .75, p = .48, η^2 = .01. These results thus indicate that the intervention group significantly improved their implicit ability beliefs only during the first research phase.

Positive effort beliefs. Repeated measures ANOVA results also revealed no significant main effect of time on pupils' Positive Effort Belief scores from T2 to T3, Λ = 1.0, $F(1, 149)$ = .40, p = .53, η^2 = .00. In addition, we did not find an interaction effect between time and grade level for this component, Λ = .98, $F(1, 149)$ = 3.30, p = .07, η^2 = .02. Thus, in contrast to *Hypotheses 5b* and *8*, no further improvement was observed in the case of pupils' positive effort beliefs during the second research phase. At T3, ANOVA results indicated no significant mean score differences among the 5th and 6th grade pupils on the Positive Belief component, $F(1, 153)$ = .26, p = .61, η^2 < .01.

Similarly, as shown in Table 5.5, repeated measures ANOVA indicated no significant main effect of time, nor an interaction effect between time and grade level, from T1 to T3 for the Positive Effort Beliefs component. Thus, the intervention group did not further improve their Positive Effort Beliefs during the second phase of the program, but on average, did sustain their improved scores from the first phase of the program.

Achievement goal orientation. Lastly, repeated measures MANOVA results only indicated a significant interaction effect between time and grade level from T1 to T3, Λ = .94, $F(2, 146)$ = 4.81, p = .01, η^2 = .06. Table 5.5 shows that pupils' Performance-Avoidance Goal Orientation mean scores in the 5th grade had significantly decreased from T2 to T3, while the mean scores of the 6th grade pupils had remained statistically unchanged. In addition, however, pupils' Mastery Goal Orientation mean scores in the 6th grade had significantly *decreased* from T2 to T3,

while no significant change was observed for the 5th grade pupils on this component. Thus, in line with *Hypotheses 6b* and *8*, the school intervention program succeeded in further improving pupils' goal orientation motivation during the second research phase, particularly in decreasing the 5th grade pupils' focus on performance-avoidance goals. MANOVA results indicated no significant mean score differences among the two grade level groups at T3, $F(2, 152) = .42$, $p = .66$, $\eta^2 = .01$.

Repeated measures MANOVA results indicated a significant main effect of time across both goal orientation components, $\Lambda = .92$, $F(2, 132) = 5.70$, $p < .01$, $\eta^2 = .08$. However, as shown in Table 5.5, pupils only significantly decreased their Performance-Avoidance Goal Orientation mean scores from T1 to T3. Furthermore, Table 5.5 shows that the 6th grade pupils significantly increased their Mastery Goal Orientation mean score and decreased their Performance-Avoidance Goal Orientation mean score from T1 to T3. In line with our predications, these results thus indicate that the school intervention program improved pupils' goal orientation motivations during the first and the second phase of the program.

Discussion

Main findings and implications

Relationships among pupils' attitudes, beliefs and motivations

The influence of attitude towards epistemic curiosity on positive effort beliefs and goal orientation motivations. This study provided empirical support for the proposition that pupils' attitudes towards epistemic curiosity predict their effort beliefs and achievement goal orientation motivation to learn by inquiry (*Hypothesis 1*, largely supported). Thus, in line with attitude and motivation theory (e.g., Ajzen, 2001; Eagly & Chaiken, 1993), pupils' attitudes towards epistemic curiosity may form important precursors to pupils' inquiry behavior in the classroom and should therefore be prioritized as a target variable in inquiry-focused school intervention programs.

In this respect, we highlight the substantial predictive power of pupils' perceptions about the personal use of epistemic curiosity (i.e., Personal Inclination) on their efforts and motivations that relate to inquiry ($\beta > .68$). This finding is in line with attitude theory, which claims that perceptions about behavioral outcomes (as part of other types of attitudinal perceptions described by the TPB) predominantly influence behavior (Ajzen, 2001). This implies that, among the five attitude components, the Personal Inclination component may function as a particularly

salient precursor to pupils' potential inquiry behavior and should therefore receive particular attention by teachers.

Contrary to our expectations, however, no statistically significant relations were found between pupils' self-efficacy and their positive effort beliefs and goal orientation motivations. In our view, this result can be explained by pupils' minimal experience and practice with curiosity-focused (inquiry) learning at the onset of the study. In fact, explorative studies suggest that curiosity-driven learning in primary school is typically scarce (see Post & Walma van der Molen, 2018a, Engel & Randall, 2009). In line with Post and Walma van der Molen (2018b), we believe that the relationship between pupils' self-efficacy and their effort beliefs and achievement goal orientations may likely become more apparent (i.e., statistically significant) as pupils gain (positive) experiences with inquiry over time. Unfortunately, our current sample size of the pupils in the intervention group proved too small to investigate these changes by means of SEM. We thus recommend future studies to examine such possible intervention effects among a larger sample of primary school pupils.

The influence of implicit ability beliefs on positive effort beliefs and goal orientation motivations. The current results also indicated that pupils' implicit beliefs about the malleability of their abilities may function as influential precursors to their effort beliefs and achievement goal orientation motivation. Thus, in line with motivation theory (e.g., Blackwell et al., 2007; Dweck, 2000), pupils' implicit ability beliefs should be made an explicit objective in teachers' inquiry-focused teaching as well (*Hypothesis 2*, supported). Furthermore, we highlight that pupils' incremental ability beliefs may function as particular salient precursors to their efforts and motivations that relate to inquiry, as the corresponding regression pathway coefficients were generally high throughout the program ($\beta > .79$).

We also note that pupils' entity beliefs positively predicted positive effort beliefs and mastery goal orientation (but to a much lesser extent than their incremental ability beliefs). This result supports the propositions by Burnette et al. (2013) that pupils who hold entity beliefs may still exert effort to learn and focus on mastery goals. However, as others have stressed as well, the potential negative influence of entity beliefs on effort beliefs and goal orientation motivations seems to occur mostly when pupils experience learning setbacks (e.g., failing for a test), which then likely leads them to disengage from and avoid learning due to their fixed ability beliefs (see also Van Aalderen-Smeets & Walma van der Molen, 2018a). For practical constraints of the study, we did not measure and include pupils' setback experiences as a separate

variable (i.e., moderator) in Model 2. Future studies should investigate this possible moderation effect.

Effects of the school intervention program on pupil growth. Results of the current study also indicated that the school intervention program positively influenced pupils' attitudes towards epistemic curiosity, implicit ability beliefs, and goal orientation motivations. At the condition level, however, these improvements proved to be less widespread than hypothesized. Pupils' implicit ability beliefs and goal orientation motivations improved as a result of the program (*Hypotheses 4a and 6a*, supported), while their attitudes and effort beliefs did not (*Hypotheses 3a and 5a*, partially rejected). At the individual grade level, the success of the intervention proved to be much more significant and widespread. While the mean scores of the 4th grade pupils had remained statistically unchanged from T1 to T2, we observed modest improvements in the mean scores of the 5th and 6th grade pupils on most of our dependent variables. This finding thus confirmed our prediction that the higher-grade pupils would generally benefit most from the school intervention program compared to the lower-grade pupils (*Hypothesis 8*, supported).

The above grade level effects were corroborated by significant differences between the mean pretest scores of the 4th, 5th and 6th grade pupils at the onset of the study (T1). In line with our hypotheses, the T1 data revealed that the attitude, belief, and motivation mean scores of the higher-grade pupils were generally less positive (i.e., lower scores on the 'positive' variables and higher scores on the 'negative' variables) compared to the lower-grade pupils (*Hypothesis 7*, supported). This result thus supports the postulations by others (e.g., Post & Walma van der Molen, 2018a, 2018b; Engel, 2015) that overly prescriptive teaching practices in the upper grades of primary education may negatively influence pupils' attitudes, beliefs and motivations towards inquiry learning. We also note that, across the three grade level groups, pupils' attitude, belief, and motivation mean scores at T1 were only modest at best (ranging between 2 and 3 on the 4-point response option scales). These moderate scores thus underscore the relevance of school interventions aimed at stimulating positive attitudes, beliefs, and motivations towards inquiry in primary school pupils.

In sum, the above results imply that teachers' improved inquiry teaching behavior indeed affected subsequent improvements in their pupils' attitudes, beliefs and motivations during the first research phase, especially for the pupils in the 5th and 6th grades. Although we would have liked to see greater gains in pupils' attitude, belief and motivation mean scores during this time, we emphasize that this pupil growth was realized by the teachers themselves within a real-world school context, which

underscores the ecological validity of the current findings. This means that, despite of the many uncontrolled school, teacher, and pupil factors that may have confused pupils' attitude, belief and motivation development during the program, the program still succeeded in affecting significant pupil growth over time. We believe that this main result is promising for two reasons. First, it suggests that a nine-months teacher professionalization course aimed at helping teachers foster pupils' attitudes towards epistemic curiosity, effort beliefs, and goal orientation motivations, can affect significant pupil growth on these measures. Second, it suggests that such professionalization treatment seems particularly fruitful for the groups of pupils who seem to hold the least positive attitudes, beliefs and motivations towards inquiry, and thus require most attention. In sum, these findings may serve the further development of school-wide teaching practice reform initiatives in primary education that international education policies on twenty-first century learning generally promotes (e.g., OECD, 2015).

During the second phase of the program, we found that the intervention group had sustained their improved mean scores from the first research phase (T2) to the end of the second phase (T3). This suggests that the school intervention program affected durable changes in pupils' attitudes, beliefs and motivations as well. In contrast to our predictions, however, we found little further *improvements* in pupils' attitudes, beliefs, and motivations during this time (*Hypotheses 3b – 6b*, partially supported). This result is surprising, as it seemed fair to assume that teachers' ongoing stimulation of their pupils' attitudes, beliefs, and motivations would be reflected by ongoing pupil growth on these measures over time. At first glance, perhaps, our results might indicate a 'ceiling effect' of pupils' attitude, belief, and motivation scores. This would partially explain the stagnant mean scores of the 4th grade pupils of the intervention group throughout the program, which were generally as high as the mean T2 and T3 scores of the 5th and 6th grade pupils. In our view, however, this possible limit of pupils' scores seems unlikely, since these scores hovered around the 2 ('I somewhat disagree') and 3 ('I somewhat agree') scale response points, and thus left room for further improvement.

Alternatively, we believe that the lack of continued pupil growth during the second phase could also be explained by shortcomings in their teachers' inquiry teaching practices. As observed by Post and Walma van der Molen (2019), most teachers directly stimulated positive attitudes, beliefs and motivations in their pupils as part of daily practice (see Table 5.1). Yet, most teachers proved less successful in offering their pupils concrete opportunities for inquiry, such as by extending pupils'

creative ideas during class with design experiments or by stimulating pupils' critical thinking through discussion. This shortcoming may have limited pupils in 'testing' their positive beliefs about inquiry against personal experiences with conducting inquiry (e.g., testing whether one's epistemic curiosity behavior indeed brings about positive learning outcomes) and, subsequently, further strengthening their positive beliefs. We base this tentative explanation on attitude theory (e.g., Ajzen, 2001; Eagly & Chaiken, 1993) and motivation theories (e.g., Blackwell et al., 2007; Dweck, 2000) that propose that such confirmatory experiences moderate attitude and motivation change (see also Post & Walma van der Molen, 2018a; Van Aalderen-Smeets & Walma van der Molen, 2018a). This would imply that, while teachers can directly improve pupils' attitudes, beliefs and motivations towards inquiry to a fair degree, such as by means of explicit reflection and discussion activities, this approach should be supplemented with opportunities for pupils to personally experience the value and use of inquiry, if we want to further improve their attitudes, beliefs, and motivations in this regard.

Limitations and future research

Our main findings are promising, but should also be considered in light of several methodological limitations. First, we are careful with generalizing our results, because our study design did not provide a true naive control group comparison. As mentioned in the Methods section, the control group participated in the same teacher training course during the second research phase as the intervention group. This was done, because the control group was not willing to commit themselves to extensive teacher and pupil measurements without receiving the benefit of training. Therefore, we could not use the group for comparison during the second research phase, which limited the interpretation of our longitudinal intervention effects. It also meant that the control group was largely self-selected and that the teachers anticipated to stimulate their pupils' inquiry at a later time during the project, which might have produced bias. This seems to be reflected by the fact that pupils' attitudes, beliefs and motivation pretest scores (T1) were somewhat more positive compared to the pretest scores of the baseline reference group at this time. In our view, these indications suggest that future studies might observe greater intervention effects when involving a truly naïve group of primary school teams.

Second, we are aware that our observed changes in pupils' attitudes, beliefs and motivations were limited to the 4th, 5th, and 6th grades. Although all the teachers of the intervention group had participated in the same training course, we focused our

program evaluation on the higher-grade level groups for practical reasons. We believe that our main findings would reflect similar changes in the attitudes, beliefs and motivations of the lower-grade pupils in our sample. However, this assumption needs to be tested. Therefore, we recommend that future studies adapt our current Likert-type survey instrument to suit the lower-grade pupils, as we believe that our survey might proof too difficult for younger children to comprehend.

Third, to test our hypotheses, we conducted numerous statistical tests on a single data set. This increased the risk of finding statistically significant intervention effects merely by chance (i.e., type 1 error) and limited the extent to which current results can be generalized to the general pupil population. Notwithstanding these limitations, we believe that our observed intervention effects among the intervention group are reliable and can be accredited to the school improvement program. We adjusted for multiple testing using Bonferroni correction. Using this correction, we observed clear attitude, belief, and motivation growth among the pupils in the intervention group on most of our variables, while we observed no such growth among the pupils in the control group. To further validate our findings, however, we recommend future studies to examine the efficacy of similar inquiry-focused school improvement programs among pupils from other primary schools in other countries. Moreover, larger-scale studies would afford greater statistical power for advanced statistical testing, such as hierarchical linear modeling (Raudenbush & Bryk, 2002), to investigate what factors at the individual teacher or school level may explain differences in pupils' attitude, beliefs, and motivation growth over time.

Fourth, we note that our current SEM test results are based on cross-sectional data and, as such, do not necessarily imply *causation* among our variables. Thus, while the current study provided empirical evidence to support theories that pupils' attitudes towards epistemic curiosity and implicit ability beliefs "predict" their effort beliefs and goal orientation motivations, this assumed causality still needs to be tested. To that end, future research should examine the effects of school interventions directed at either improving pupils' attitudes towards epistemic curiosity or their implicit ability beliefs on subsequent changes in pupils' effort beliefs and goal orientation motivations, and if so, at what stages of the intervention. In our view, such insight into the causal mechanisms underlying pupils' attitude, belief, and motivation development may inform the further development of more refined school interventions aimed at stimulating pupils' inquiry thinking.

Lastly, we are well aware that the absence of behavioral data in the present study can be seen as a limitation. To our knowledge, no validated measurement instruments

yet exist to measure the wide range of complex overt and covert behaviors associated with inquiry (for details about this complexity, please see Sinatra, Heddy & Lombardi, 2015). For example, pupils may engage in inquiry not only in terms of observable behavior in the classroom (e.g., verbalized question-asking or explanation-seeking), but as much so in terms of inquiry-driven thought, such as when studying school subject matter (see also Claxton, 2007; Engel, 2006). In this study we focused on pupils' attitudes, beliefs and motivations because these aspects are considered important indicators of pupils' inquiry behavior and can be efficiently measured by means of self-report questionnaires. However, we do recommend future research into the concurrent validity of such scales, for example, by comparing pupils' attitudes, beliefs, and motivation scores with observations of their inquiry behavior in the classroom to see whether these different measures converge to similar conclusions. These comparisons should also include in-depth interviews with pupils to examine possible changes in their inquiry thought-processes. Future directions of research should thus include the development of such multi-method measures of pupils' inquiry behavior.

Conclusion

The current study shows that pupils' attitudes towards epistemic curiosity and their implicit ability beliefs likely *affect* their efforts and motivations to be inquiry-minded learners. With this study, we thus provide novel insight into the psychological mechanism that may underlie pupils' potential inquiry behavior in the classroom. This result is in line with a growing body of social psychological research that indicates that attitudinal beliefs (e.g., Ajzen, 2001) and motivational beliefs (e.g., Pintrich, 2003) determine pupils' approaches to learning and the amount of effort they put into their learning (see also Muis, Psaradellis, Lajoie, Di Leo, & Chevrier, 2015; Pekrun, 2006). The stimulation of positive attitudes towards epistemic curiosity and implicit ability beliefs may thus be a valuable supplement to typical forms of inquiry-based pedagogy.

The above finding has value for practice as well, since our current results suggest that pupils' pre-existing attitudes towards epistemic curiosity, implicit ability beliefs, effort beliefs, and achievement goal orientations in the upper grades of primary school are moderate at best and seem to worsen as pupils progress through primary school (see also Post & Walma van der Molen, 2018b; Engel, 2015). Thus, if we want to be successful in helping pupils mature into inquisitive, creative, and confident thinkers,

such as promoted by twenty-first century education standards (e.g., OECD, 2015), we advocate that teachers attend to fostering these aspects in their pupils as part of their practices.

To that end, the current study provides several practical suggestions. At its core, it suggests that teachers should make pupils aware of *why* 'being inquiry-minded' is fruitful in the first place. Teachers should recognize that pupils' attitude, belief, and motivation development in this regard is not a natural "by-product" of 'doing' inquiry, but that this development requires their explicit and daily attention (see also Post & Walma van der Molen, 2018a; Claxton, 2007; Engel, 2011; Hodson, 2014). This means that, in the midst of literacy, history, mathematics and science lessons, teachers should explicitly cultivate pupils' positive perceptions about the educational value of asking epistemic questions, a sense of pleasure and pride from seeking creative solutions to problems, and the confidence that their repeated attempts to face and tackle such problems boosts their potential inquiry ability. Through reflective group discussions, for example, teachers can make pupils aware of their (implicit) beliefs about epistemic curiosity and inquiry ability and, subsequently, challenge them to adopt more positive beliefs about these matters (for examples of such activities, see Abd-El-Khalick, 2012; Cook, Goodman, & Schulz, 2011; Deng, Chen, & Tsai, 2011). Such metacognitive lesson activities can encourage pupils to re-orient their perspectives on epistemic curiosity and on their potential inquiry ability and, consequently, motivate them to engage in inquiry more often and to further build-up their inquiry skills.

The current teacher enhancement course proved largely successful in helping teachers affect the above-described attitude, belief, and motivation change among their pupils. Please see Table 5.1 for concrete examples of the lesson interventions implemented by the teachers in this study that proved effective in this regard (for extended guidelines on stimulating pupils' attitudes towards curiosity, please see Post & Walma van der Molen, 2018a). We thus believe that the current findings provide a starting point for the further development of school development projects aimed at stimulating pupils' inquiry in primary education, namely by fostering pupils' attitudes towards epistemic curiosity and their implicit ability beliefs. For example, future studies may look into the efficacy of a shortened inquiry-focused teacher enhancement course that prioritizes on helping participants foster pupils' attitudes towards epistemic curiosity and their implicit ability beliefs. Such a shortened and more focused course would likely appeal to most primary school teams, as primary teachers generally have limited time and attention available for reforming their

practices. Furthermore, the current findings suggest that pupils' perceptions about the value of epistemic curiosity for their own learning (Personal Inclination) and about the malleability of their intellectual ability (Incremental Beliefs) may function as particularly salient precursors to pupils' potential inquiry behavior. Thus, inquiry-focused teacher professionalization could perhaps even be further focused on helping teachers stimulate and assess only these attitudinal and motivational beliefs.

At the same time, however, our findings highlight the challenge that primary teachers may face when aiming for more long-term pupil growth. As observed by Post and Walma van der Molen (in press), many teachers struggled with stimulating pupils' creative, higher-order thinking during the program. In our view, this limitation may have prevented pupils to test and further improve their attitudes, beliefs, and motivations towards inquiry. We thus underscore that teachers should *combine* an explicit focus on stimulating positive attitudes, beliefs, and motivations in pupils with concrete opportunities for pupils to conduct inquiry. According to Post and Walma van der Molen (in press), the mastery of such complex inquiry teaching skills is difficult and requires time, as teachers must also develop positive attitudes towards (teaching) inquiry and be supported in their practice reform by their school management. We expect that teachers may thus benefit from follow-up coaching in this regard. The results of the current study may provide a foundation to further investigate the details and efficacy of such integrated, long-term approaches to fostering pupils' inquiry in primary school.

Chapter 6

6　Discussion

Focus of the study

This dissertation describes the results of a two-year school improvement program, in which the complete school staffs of six Dutch primary school teams were trained to adopt inquiry-based pedagogy. International documents on primary science and technology (S&T) education increasingly promote the implementation of inquiry-based pedagogy (Lucas, Claxton, & Spencer, 2013; OECD, 2015; Osborne & Dillon, 2008). However, little is known about what it takes to help primary schools adopt inquiry-based pedagogy and to what degree an 'inquiry-infused' primary school curriculum may foster pupils' inquiry over time (Syer, Chichekian, Shore, & Aulls, 2012).

Therefore, in this dissertation, the aforementioned program was developed and evaluated. The program was unique in three ways. First, it adhered to a broad focus on the meaning of 'S&T education' as recently promoted in the literature, in which pupils not only learned *about* inquiry or *to* inquire, but *used* inquiry as a general strategy to study school subject matter. This focus is believed to provide teachers increased opportunity to meaningfully and structurally incorporate S&T teaching in their regular education programs (Heywood et al., 2012; Osborne, 2014). Second, the program included *substantial* teacher professionalization aimed at helping participants acquire the competency to teach (S&T) by inquiry as part of their daily practices, rather than professionalization aimed at increasing teachers' S&T content knowledge or for the implementation of prescribed S&T lessons (see Slavin, Lake, Hanley, & Thurston, 2014; Van Aalderen-Smeets & Walma van der Molen, 2015). Third, the program adopted a *school-wide* approach to the implementation of inquiry-based pedagogy, rather than involving only few individual school members. This was decided, because teachers' individual professional development is believed to be dependent on school leadership and organization, such as school policy, teacher collaboration, and school culture (Moolenaar, Sleegers, & Daly, 2012; Uiterwijk-Luijk, Krüger, & Volman, 2019). Therefore, we encouraged all school principals to act as inquiry-oriented leaders of teachers' practice reform during the program.

Given the above, the program effects were examined on multiple school levels, at multiple measurement times, and by multiple measures to gain a thorough understanding of the efficacy of the school improvement program. It was expected

that teachers' attitudes towards inquiry teaching and their inquiry teaching behavior would improve as a result of participating in the program. We also examined changes in school leadership during the program and how variation in leadership could explain variance in the effects of the program on teachers' professional development. At the pupil level, we expected changes in teachers' inquiry teaching practices to affect subsequent changes in pupils' inquiry-related attitudes, beliefs and motivations during the program.

In the last two studies of this dissertation (chapters 4 and 5), the above expectations were tested. In the first two studies of this dissertation (chapters 2 and 3), we investigated pupils' curiosity in primary school. Few scientific guidelines and instruments previously existed to encourage and assess pupils' curiosity within school settings (see Grossnickle, 2016). This shortcoming was problematic, because the encouragement of pupils' curiosity is widely promoted as one of the main objectives of S&T teaching in primary education (e.g., Carr & Claxton, 2004; Jirout & Klahr, 2012; Verkenningscommissie, 2013). In the development of the current school improvement program, we thus needed to develop these guidelines and measures first. Based on attitude and curiosity research, we propose that pupils' curiosity can be understood in terms of their *attitudes* towards curiosity and that, therefore, its positive development should be regarded as a primary objective of inquiry-based pedagogy. This attitudinal perspective on pupils' curiosity is closely related to the proposed definition of 'S&T-minded' pupils by the Dutch Exploratory Committee (Verkenningscommissie, 2013), as described in the first chapter of this dissertation. We thus adopted this objective in the development and evaluation of our program.

In sum, this dissertation contributes to the knowledge base of primary science and technology education by describing the features and effects of a comprehensive two-year inquiry-focused school intervention. This final chapter provides an overview of the most important findings of the research presented in this dissertation. This overview is followed by a discussion of the potential limitations of the overall study. Finally, implications and recommendations for future research and practice are presented.

Important findings of the study

Investigating pupils' own curiosity beliefs, feelings and experiences

As explained in the first chapter of this dissertation, we were somewhat surprised to discover a scarcity of research on pupils' curiosity within educational settings (see also

Grossnickle, 2016). Curiosity can be understood as the desire to seek and obtain new intellectual information (Piotrowski, Litman, & Valkenburg, 2014). The few available studies on this topic have generally been limited to qualitative, explorative research (e.g., Engel, 2015) and suggest that pupils are provided little opportunity and encouragement to be curious in primary school. Researchers and policy-makers have often attempted to define curiosity *for* pupils (e.g., Jirout & Klahr, 2012), but we did not come across any studies that investigated pupils' *own* concepts about, feelings towards, and experiences with 'being curious' in school. Such research, however, brings insight into the aspects of pupils' curiosity that may require improvement and, as such, serve as an important stepping stone for the development of curiosity-focused pedagogy, teacher professionalization, and measurement instruments (e.g., Bransford, Brown, & Cocking, 2000)

In this dissertation (chapter 2), we thus first set out to explore primary pupils' pre-existing concepts of, feelings towards and experiences with 'curiosity' from all grade levels of two Dutch primary schools. The results of the study indicated that most pupils barely understood curiosity in epistemic terms, but that they associated the term predominantly with something that belongs to the social domain (e.g., eavesdropping, prying, gossiping). Partly because of this narrow conception, the pupils did not attribute much learning-value to being curious in school. In addition, they felt generally discouraged by their teachers to express their epistemic questions and ideas in class. However, many pupils reported to be actively curious about a diverse range of complex science topics outside of the school context, such as at home or during holiday trips. This result is in line with studies in the curiosity literature that pupils are naturally curious (e.g., Warren, Ballenger, Ogonowski, Rosebery, & Hudicourt-Barnes, 2001).

As part of the debriefing of the study, most of pupils' teachers in the study indicated to feel concerned about pupils' limited curiosity experiences in school, as they all perceived the clear educational value of pupils' curiosity. Many teachers indicated to perceive pupils' curiosity as a natural "by-product" of their inquiry-based lesson activities and that its stimulation did thus not require their explicit attention. The results of this study suggest that this perception is likely false and that, therefore, teachers *should* explicitly foster pupils' epistemic curiosity as part of their regular lesson objectives (see also Claxton, 2007; Hodson, 2014). In other words, no matter how much we try to stimulate pupils' curiosity with attractive and diverse curiosity-eliciting lesson activities, if we do not make them aware of *why* we prompt them to be curious, they might just consider these attempts as meaningless behaviour and

remain reluctant to curiously engage in other learning. This approach to stimulating pupils' inquiry is in line with the propositions by others in this field (e.g., Akerson, Abd-El- Khalick, & Lederman, 2000; Deng et al., 2011).

Interpreting pupils' curiosity in terms of attitude
The above findings led us to consider the aspects of pupils' curiosity that would require particular attention by primary teachers and, as such, determine the focal point of related teacher professional development, educational activities, and measures. These considerations are described as part of the second study of this dissertation (chapter 3).

Rather than focusing on pupils' curiosity behavior, skills, traits, or states, we propose to shift focus towards the positive development and assessment of pupils' *perceptions* of curiosity. Such perceptions can be understood in terms of pupils' images of and attitudes towards curiosity (for example, see Ajzen, 2001). A large body of research shows that attitudinal and motivational beliefs affect behavior (e.g., Ajzen & Fishbein, 1980; Eagly & Chaiken, 1993), such as the types of emotions pupils experience during learning (Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011) and the types of learning strategies they employ (Muis, Psaradellis, Lajoie, Di Leo, & Chevrier, 2015). Based on attitude and curiosity theory, we thus advocate that a fruitful approach to stimulating pupils' epistemic wonderment, questions, and ideas in the classroom is to foster their epistemic images of curiosity and their positive beliefs and feelings about being inquisitive learners. As mentioned before, this attitudinal perspective on the stimulation of pupils' curiosity closely matched the proposed definition of 'S&T-minded' pupils by the Exploratory Committee (Verkenningscommissie, 2013).

Shifting focus towards the investigation of pupils' images of and attitudes towards curiosity also provides practical advantages, as attitude can be reliably measured and positively developed (Vogel & Wänke, 2016). Notably, most curiosity definitions in the scientific literature have so far been critiqued for confounding the concept of curiosity with the concepts of interest, intelligence, or motivation, which makes it unclear what is measured or what mechanisms may underlie pupils' curiosity behavior (Grossnickle, 2016; Silvia & Sanders, 2010). Our newly proposed attitudinal perspective on pupils' curiosity thus proved useful in these respects. However, because this perspective was new, no corresponding operationalizations or measurement instrument yet existed. As part of the second study (chapter 3), we thus

made use of attitude, motivation, and curiosity theory to propose seven components of pupils' images of and attitudes towards curiosity.

We distinguished pupils' social and epistemic images of curiosity as two separate components of pupils' images of curiosity. Pupils' social images of curiosity represent behaviors of seeking or obtaining new information about social experiences, such as eavesdropping and prying (Litman, Robinson, & Demetre, 2016). Pupils' epistemic images of curiosity represent behaviors of seeking or obtaining new intellectual information, such as asking follow-up questions to explore subject matter (Piotrowski et al., 2014). Based on the underlying dimensions of the Theory of Planned Behavior by Ajzen (2001), we distinguished five separate components of pupils' attitudes towards epistemic curiosity: pupils' perception of the learning value and pleasure of being epistemically curious (Personal Inclination); pupils' perception of the value of epistemically curious people to society (Societal Relevance); pupils' perception of their classmates' negative judgments about them being curious (Fear of Classmates' Negative Judgment); pupils' negative opinions about other people's epistemic curiosity (Negative Opinion); and lastly, pupils' perception of their capability to be epistemically curious (Self-Efficacy).

In sum, the second study of this dissertation thus contributes to curiosity research by unpacking the overarching concept of pupils' curiosity in its underlying image and attitude components. It thereby, also, offers a novel and relevant perspective on the stimulation and assessment of pupils' curiosity in primary school, namely by positively developing pupils' images of and attitudes towards epistemic curiosity. To that end, primary teachers should foster positive classroom cultures in which pupils (learn to) perceive the educational value of being inquisitive learners in school, derive a sense of pleasure and pride from asking inquisitive questions or wondering about alternative answers to questions, and feel that their teachers and peers encourage such behavior.

Instrument development and validation

As part of the second study, we translated the above-described image and attitude components of pupils' curiosity into corresponding Likert-type measurement scales, which made up the Children's Images of and Attitudes towards Curiosity (CIAC) questionnaire. This translation enabled us to empirically examine the construct validity of the scales and, in addition, allowed us to measure (changes in) pupils' images of and attitudes towards epistemic curiosity as part of our planned inquiry-focused school intervention in studies 3 and 4 (chapters 4 and 5).

Because the CIAC questionnaire and its underlying theoretical components had not been empirically tested before, we used qualitative and extensive quantitative methods to verify its construct validity. Although factor analytic results showed some minor deviations from the constructs that we originally hypothesized, results supported the construct validity of the CIAC. Results also indicated that the CIAC measures the same image and attitude components the same way for pupils across the 4th, 5th and 6th grades. This makes the CIAC useful for assessing longitudinal changes in pupils' images and attitudes. Consistent with attitude and curiosity theory (e.g., Eagly & Chaiken, 1993; Grossnickle, 2016), we also observed positive relations between pupils' images of curiosity, attitudes towards epistemic curiosity, and their achievement goal orientation motivations, which provided support for the predictive validity of the CIAC as well.

In addition, as described in the introduction of this dissertation, we aimed to assess a broad range of pupil, teacher and school principal variables to evaluate the efficacy of our professionalization treatment. The development of these variables and their corresponding scales is not described in separate chapters of this dissertation, but included on a smaller scale as part of the third and fourth study (chapters 4 and 5). In our view, instrument validation is highly important in educational science (see also Shore, Aulls, & Delcourt, 2017), especially in the domain of inquiry-based teaching and learning, since few validated measurement instruments exist and, when available, are often critiqued for showing poor psychometric reliability or theoretical overlap with other measures (e.g., Carr & Claxton, 2004; Pellegrino & Hilton, 2012). As part of this dissertation, we thus set out to develop and validate (new) relevant outcome variables of inquiry-focused school development.

To that end, at the pupil level, we developed the CIAC questionnaire to measure (changes in) pupils' attitudes towards epistemic curiosity during the school improvement program. In addition, we adapted existing measurement scales for implicit ability beliefs (i.e., perceived malleability of their ability), effort beliefs (i.e., perceived causality of their effort on achievement), and achievement goal orientations (i.e., perceived goals for achieving in school) to suit our target group and validated these scales by factor analyses. In the last study of this dissertation, we show that pupils' attitudes towards epistemic curiosity and their implicit ability beliefs significantly affect their efforts and motivations to be inquiry-driven learners. This finding is in line with attitude and motivation theory (e.g., Ajzen & Fishbein, 1980; Pintrich, 2003; Blackwell, Trzesniewski, & Dweck, 2007) and thereby provides (further) support for the predictive validity of our scales.

At the teacher level, teachers' attitudes towards inquiry teaching were assessed by means of the DAS questionnaire that was previously developed and validated by our research group (see Van Aalderen-Smeets & Walma van der Molen, 2013). No measures previously existed to measure teachers' perceptions of creative (inquiry-enhanced) lesson design. We thus newly developed a corresponding measurement scale. Also, no validated instruments were available yet to assess teachers' inquiry teaching behavior, such as aimed at stimulating pupils' attitudes towards curiosity, implicit ability beliefs, and their higher-order thinking. Therefore, we devised a method to gather perception data from individual teachers and groups of their pupils on teachers' inquiry behavior. By cross-examining pupils' and teachers' perception scores, we developed an integrated understanding of teachers' behavior development during the program (see also Van der Scheer, Bijlsma, & Glas, 2018). This method provided a reliable and efficient way of assessing teachers' behavior development during the program. Similarly, we gathered perception data from the teachers and their school principals to measure (changes in) school principals' leadership on teachers' adoption of inquiry-based pedagogy during the program, such as in terms of their leadership on school policy, teacher collaboration and classroom experimentation, and school culture.

Given the above work, this dissertation contributes to inquiry research by offering relevant and validated measurement instruments for evaluating inquiry-focused school interventions. These instruments can be used by other researchers to examine the effects of similar interventions.

School-wide teacher professionalization on inquiry teaching
In the development of the present professionalization treatment, we attempted to meet the above-described target variables of inquiry-focused school development. Thus, we developed a comprehensive inquiry-focused school intervention, which aimed at positively developing the knowledge, skills and attitudes of the complete school staffs of six Dutch primary schools for integrating inquiry-based pedagogy into their entire school organization. To that end, the intervention included the complete six-months attitude-focused teacher professional development course by Van Aalderen-Smeets and Walma van der Molen (2015). In addition, we developed a three-month training course that aimed to help participants familiarize with the broader concept of S&T teaching and learning (i.e., fostering pupils to learn *by* inquiry). It also aimed at helping participants acquire the didactic knowledge, skills, and positive attitudes to foster pupils' inquiry as part of their daily practices. The

participating school leaders did not receive specialized leadership training, but were required to participate in all course meetings and were encouraged to facilitate teachers' adoption of inquiry-based pedagogy during the program. For details on the contents of the course, please see chapter 4 of this dissertation.

As described in chapters 4 and 5, our analyses indicated that the professional training significantly improved teachers' attitudes and inquiry teaching behavior, and that teachers' improved practices persisted to one year after the training. Teachers' improved inquiry teaching practices affected significant improvements in pupils' attitudes, beliefs and motivations over time. However, school leadership remained largely moderate throughout the program and appeared to explain little variation in teachers' inquiry teaching development among the participating schools. Also, the same training showed a limited effect in the case of the delayed intervention group, which seemed to be caused by incidental school-specific factors that hindered school improvement (e.g., mismatch of training needs, being discontent about the personal style of one of the trainers, the introduction of new lesson methods). In sum, we thus feel safe to say that the program proved largely successful in helping primary school teams integrate inquiry-based pedagogy into their school organizations and, in effect, fostered pupils' inquiry-related attitudes, beliefs, and motivations. Many of the participating teachers and school principals stated at the end of the research project that the program met their expectations and professionalization needs.

We note that results also indicated that the program affected somewhat different changes in teachers' attitude and behavior development among the participating schools. In line with the school improvement literature (e.g., Berliner, 2002), this finding indicates the importance of examining school (organization) characteristics before, during, and after professionalization treatment, and to investigate how variation in the development in these characteristics over time may explain variance in the effects of professionalization treatment on teachers' professional development.

Lastly, we note the importance of 'school-wide' approaches to the implementation of inquiry-based pedagogy. Results suggest that teachers' attitude development was significantly fostered by teachers' collective reflections and discussions about their adoption of inquiry-based pedagogy. However, in the present study, the school leaders did not seem to play a significant role in further fostering these activities during the program. According to most of the teachers, their school principals failed to set clear expectations on their inquiry teaching practice reform, did not actively facilitate teacher collaboration on this front, and provided little

encouragement for classroom experimentation. This result is disappointing, as it is believed that school principals can play an important role in fostering school practice reform in these ways (for example, see Leithwood, Harris, & Hopkins, 2008). Apparently, the present program did thus not succeed in enabling school principals to act as inquiry-oriented leaders. Regardless, most teachers still managed to reform their practices and proved successful in positively developing their pupils' inquiry-related attitudes, beliefs, and motivation during the two years of the program. Therefore, the question rises whether more explicit attention for the development of school leadership on the implementation of inquiry-based pedagogy would actually make a significant difference in teachers' inquiry teaching development.

Longitudinal investigation of pupils' attitudes, beliefs and motivations
The above results are promising and underscore the vital role that primary teachers can play in fostering pupils' inquiry learning in school as an integral component of their regular education programs. At the same time, these results indicate that helping primary school teams adopt inquiry-based pedagogy is a complex enterprise, requiring substantial professionalization and at least several years for implementation (see also Capps, Crawford, & Constas, 2012; Desimone, 2009). Such extensive professionalization is particularly relevant for realizing significant attitude change among participants, because such change naturally requires time for participants to reflect on, familiarize and gain positive experiences with inquiry teaching (Van Aalderen-Smeets & Walma van der Molen, 2015). Despite this substantial and longitudinal approach to helping the school teams adopt inquiry-based pedagogy, however, we note that the program still did not affect significant changes on *all* of our outcome variables. We thus underscore the importance of substantial and long-term teacher professionalization to help primary schools fully adopt inquiry-based pedagogy (see also Walan et al., 2016).

In addition, we show in chapter 5 that teachers' improved inquiry teaching behavior affected significant improvements in pupils' attitudes, beliefs and motivations during the first year of the program. However, pupils' development did not continue during the second year of the program. In particular, many teachers indicated to struggle with stimulating pupils' creative, higher-order thinking during lesson activities (as opposed to designing inquiry-eliciting study assignments or research projects for pupils to work on). In our view, this shortcoming may have prevented pupils to test their positive beliefs about inquiry against actual positive experiences of 'doing inquiry' during class (e.g., perceiving that pursuing their

curiosity actually leads to improved knowledge attainment), which consequently prevented them to further improve their attitudes, beliefs, and motivations towards inquiry during the second year of the program. This finding indicates that the positive development of pupils' attitudes, beliefs, and motivations by *direct* stimulation is limited. To advance pupils' development in this regard, teachers should thus combine an explicit focus on stimulating pupils' positive attitudes, beliefs, and motivations with opportunities for them to conduct inquiry. We expect that teachers may benefit from follow-up coaching in stimulating pupils' higher-order thinking during lesson activities.

In sum, the above results thus indicate the importance of longitudinal investigations into pupils' development during inquiry-focused school interventions.

Potential limitations of the study

Limited selection of schools

The above-described findings and implications of this study are promising, but should also be interpreted in light of several potential limitations. First, we note that six whole primary school teams (including *all* school staff) participated in a nine-months inquiry-focused teacher professionalization course. Although such large-scale intervention research is rare, larger-scale research would have likely afforded us with a greater and more diverse pupil, teacher, and school principal sample. This would have provided greater insight into the possible variation of school organization practices among primary schools and its influence on the effects of the professionalization treatment on teachers' and pupils' development. In addition, larger-scale data collection would have also improved the statistical power of our analyses and allowed for more advanced (hierarchical) statistical testing (Raudenbush & Bryk, 2002). Notwithstanding these limitations, the current scale of the project proved to be maximally attainable as part of a single doctoral research project. We also note that our current research design and methodology still offered sufficient opportunity to experimentally test the efficacy of the current school improvement program.

We are also careful with generalizing the findings of the study. We enrolled six primary schools to take part in the present program based on their interests in the goals of the program and on their commitment to take part in our extensive measurements. We also required all schools to pay for participation. While selecting highly committed schools for participation was deemed important for practical

reasons, this means that the positive effects found in the present study may apply predominantly to *committed* primary school teams. Less committed schools might have lacked the motivation to complete the professionalization treatment. Alternatively, we might have found greater program effects among such school teams, as teachers of these schools would likely show less positive pretest scores compared to teachers from committed school teams, and thereby show increased improvement as a result of participating in the program. In fact, as described in study 4 (chapter 5), pupils' mean pretest scores of the baseline reference group (i.e., the three primary schools that did not take part in the intervention) were slightly, but statistically, less positive compared to the mean pretest scores of the pupils in the treatment schools on most of our variables. However, as part of study 3 (chapter 4), no statistical differences were found in teachers' attitude mean scores at the time of the pretest between any of the schools, including the baseline reference group.

Irrespectively, we note that all six schools had little experience with the broader 'type' of inquiry-focused (S&T) education as promoted in this dissertation. We thus believe that all of the treatment schools were representative of the general population of primary schools in The Netherlands.

Lack of a true control comparison group

We also note that our delayed treatment pretest-posttest control group design did not provide a true naive control group comparison during both phases of the intervention. During the first phase, the delayed treatment group was already aware of the goals of the program and anticipated inquiry-focused professional training the following school year. During the second research phase, both treatment groups (had) received training and, therefore, we needed to rely on within-group comparisons of participants' scores to assess school development.

Recruiting school teams for control in any longitudinal school development research project is difficult. School teams that do not perceive the relevance of such a project are not likely to commit themselves to multiple and extensive measurements without receiving the benefit of training. This is unfortunate, but understandable. Our delayed treatment control group design thus proved useful in this respect. However, because of this decision, we do not know for certain whether our observed changes in pupils', teachers' and school principals' scores during the program can be attributed to the intervention or whether these effects are affected by other unobserved factors. We feel safe to say that the intervention was effective, both in the first and second phase of the research, because our quantitative results were corroborated by

qualitative interview data collected from many of the pupils, teachers and school principals of each treatment school. In this regard, we underscore the importance of mixed-method approaches to data collection and the assessment of (inquiry-focused) school development.

No behavioral measures of pupils' inquiry

As previously described in this chapter, we evaluated the effects of the program on teachers' and school principals' behavior on the basis of perception data. This method appeared to produce reliable patterns in respondents' perception scores and indicated that the teachers significantly improved their inquiry teaching behavior during the program.

In this dissertation, however, we did not collect data on *pupils'* inquiry behavior development during the program. Instead, we assessed pupils' inquiry-related attitude, belief and motivation development as relevant indicators of their potential inquiry behavior. We are well aware that the investigation of pupils' inquiry behavior would have been relevant, as this would have enabled us to empirically test our assumed attitude-behavior and motivation-behavior relations among the pupils and, in addition, allowed us to assess the extent to which the program positively affected behavioral changes in pupils (e.g., inquisitive question-asking, proposing creative ideas for solving complex problems, connecting school subject matter).

However, the measurement of pupils' inquiry behavior is complex, as pupils may engage in inquiry not only in terms of observable behavior in the classroom (e.g., verbalized question-asking or explanation-seeking), but as much so in terms of inquiry-driven thought (see also Claxton, 2007; Engel, 2015; Sinatra, Heddy & Lombardi, 2015). We focused on pupils' attitudes, beliefs and motivations because these aspects are considered important indicators of pupils' inquiry behavior and can be reliably and efficiently measured by means of self-report questionnaires. We recommend future research to look into the concurrent validity of our scales, for example, by comparing pupils' attitudes, beliefs, and motivation scores with observations of their inquiry behavior in the classroom to see whether these different measures converge to similar conclusions.

Alternatively, we have considered examining possible changes in pupils' scores on Dutch national knowledge tests as indications of pupils' developing inquiry competency during the program. For example, the Cito exams (in Dutch: Cito-toets) comprise a total of 290 multiple-choice questions, which test pupils' Dutch reading comprehension skills, mathematics, world orientation (i.e., geography, biology, and

history) and study skills. One could imagine positive effects of the program on these measures, especially in the cases of world orientation and study skills, as inquiry-based pedagogy is believed to generally foster pupils' strategies to investigate the interrelatedness of different subject domains (Bruner, 1961; Papert, 1980; Syer et al., 2012). However, upon review of the test items, we felt that few items 'tested' the kind of inquiry skills and knowledge attainment promoted in this dissertation. At most, items required pupils to interpret graphs or derive relevant information from a piece of text. In addition, we found that only few of the schools that participated in the present study administered the (same) Cito exams. These shortcomings thus prevented us from pursuing this direction of analysis.

Recommendations for future research

Development of measures of inquiry-based pedagogy and school development

As promoted in this dissertation, there is increasing consensus in the school improvement literature about the importance of *school-wide* implementations of inquiry-based pedagogy (e.g., Syer et al., 2012). Multi-level approaches to the development and evaluation of such implementations have so far been scarce (Thurlings et al., 2015). Moreover, few measurement instruments exist to assess inquiry-focused school development and, when available, are often critiqued for showing poor psychometric quality (e.g., Pellegrino & Hilton, 2012).

With the present study, we derived relevant outcome variables from the relevant literatures not only at the level of the *teacher* (e.g., attitudes towards inquiry teaching, creative and prescribed lesson design perceptions, inquiry teaching behavior), but also at the level of the *school principal* (e.g., leadership behavior focused on school policy, inquiry teaching support, school culture) and the *pupils* (e.g., inquiry-related attitudes, beliefs and motivations). In addition, we developed and validated corresponding measurement instruments. These instruments may be used by other researchers to reliably and efficiently assess the effectiveness of similar inquiry-focused school interventions.

As part of this dissertation, however, we advocate the importance of further research into the development of other relevant outcome variables of inquiry-focused school development. These variables may serve as guides for the research field of what variables are essential to foster and allow comparison of results across studies to further build our knowledge base. Corresponding measurement instruments can later

be used by schools and trainers for summative and formative assessment (see also Geisinger, 2016).

Potential use of pupil perceptions to evaluate inquiry-based pedagogy
At the pupil level, we recommend researchers to make use of our Likert-type questionnaires to assess (changes in) pupils' inquiry-related attitudes, beliefs and motivations as indicators of their potential inquiry behavior. However, as discussed in this chapter, we are well aware of the need of measures of pupils' inquiry learning behavior. The findings of the present study might provide a fruitful direction for future research on this front.

Most upper-grade pupils in our sample were able to differentiate between various complex constructs of their teachers' inquiry teaching behavior. This finding suggest that upper-grade primary school pupils might be able to rate complex aspects of their *own* (overt and covert) inquiry learning behavior in the classroom as well. Combined with teachers' perceptions about pupils' (overt) behavioral development, pupils' and teachers' perceptions can be cross-examined to efficiently and reliably assess pupils' inquiry behavior development during interventions.

Future research should examine whether this approach produces valid results. To that end, variation in pupils' individual perception scores could be compared to the variation in their teachers' perception scores, and pupils' own attitude, belief, and motivation scores.

Comparison of different inquiry-focused school interventions
Lastly, we recommend future research to compare the effects of the present intervention with other professional development programs. One could argue that it is not surprising that training has more effect than no training, as was found in the present study. It might thus be valuable to compare the effects of our current program to a group of schools that received traditional inquiry-focused professionalization (e.g., implementation of prescribed inquiry lessons or attitude training only).

Such comparative studies may bring insight into the contribution of particular features of inquiry-focused school improvement programs to teachers' potential inquiry teaching development. Such insight may also be used to offer schools more targeted professionalization trajectories. The latter may be especially relevant, as individual schools differ in terms of their team member constellations, prior experiences with inquiry teaching, and school improvement goals. As advocated throughout this dissertation, we therefore underscore the importance of examining

the effects of inquiry-focused school interventions both at the treatment level and at the individual school level.

Recommendations for practice

Focus on pupils' attitudes, beliefs, and motivations

One of the main findings of this dissertation is the importance of stimulating pupils' attitudes towards epistemic curiosity and their implicit beliefs about the malleability of their inquiry ability. While cognitive and hands-on inquiry-based lesson activities remain important to build-up pupils' inquiry knowledge and skills, our findings suggest that such activities do not necessarily foster pupils' attitudes, beliefs and motivations that relate to inquiry. This finding is in line with other studies, which indicate that the relation between increased inquiry knowledge skills and skills on the one hand and inquiry-related attitudes on the other remains unclear (Avery & Meyer, 2012; Ucar & Demircioglu, 2011).

As discussed in study 1 (chapter 2), inquiry teaching should therefore incorporate *metacognitive* activities that stimulate pupils to consider their existing concepts, feelings and experiences about 'curiosity' and their 'inquiry ability', to compare these with more epistemic and incremental notions, to discuss the possible uses of curiosity and effortful inquiry for their own learning, and to learn to recognize and articulate their own wonderments, questions and ideas in more epistemic and incremental terms.

Such metacognitive activities can be held either as stand-alone lesson interventions or held in relation to particular subject matter or research projects (see Deng et al., 2011). In parallel, teachers may guide such group discussions by illustrating the ways that epistemic curiosity drives pupils (and professionals) to explore new subject matter, to creatively come to new ideas or alternative explanations to existing problems, and to improve their overall academic achievement. They should also become 'responsive' to pupils' emergent comments, ideas and questions during lesson activities as starting points of metacognitive discussion. For example, teachers can emphasize the educational value of asking inquisitive questions about subject matter when a pupil asks such a question during their study of a school topic. Similarly, pupils' negative remarks about the curiosity behavior of their peers may serve as meaningful situation to discuss pupils' images of curiosity and, consequently, inspire them to adopt more epistemic images of inquiry and positive beliefs about epistemic curiosity behavior in the classroom.

At the same time, findings of study 4 (chapter 5) suggest that the above kind of direct reflection and discussion is not sufficient to fully develop pupils' attitudes, beliefs and motivations. Such activities should be supplemented with concrete opportunities for pupils to *experience* the personal value and use of inquiry for learning as well, such as by perceiving and joy of studying novel and complex subject matter, the challenge of working on open-ended and ill-structured research projects, and the value of employing their higher-order thinking to come up with original and creative ideas (e.g., Anderson & Krathwohl, 2001; King, Goodson, & Rohani, 2011; Lederman & Abell, 2014). This allows pupils to 'test' their positive beliefs about inquiry against personal experiences with conducting inquiry (e.g., testing whether one's epistemic curiosity behavior indeed brings about positive learning outcomes) and, subsequently, may further strengthen their positive beliefs. We thus recommend that primary teachers combine an explicit focus on stimulating positive attitudes, beliefs, and motivations in pupils with sufficient opportunity for pupils to conduct inquiry.

Features of effective inquiry-focused teacher professionalization
With the present study, we show that our nine-months teacher professional development course positively affected teachers' attitudes towards inquiry teaching and their inquiry teaching behavior, and that teachers' improved practices persisted to one year after training. Based on these results, we feel safe to say that the features of the present program provide a foundation for the further development of inquiry-focused school interventions.

First, our results support theories in the literature that a combination of extensive attitude-focused and didactical training provides primary school teams the (minimal) preparatory training to implement and sustain inquiry-focused teaching practice (e.g., Osborne & Dillon, 2008; Van Aalderen-Smeets & Walma van der Molen, 2015). We highlight that school teams, even those who are highly committed to adopt such practice, do not necessarily possess positive attitudes towards inquiry teaching to begin with. As shown in the third study, teachers' self-efficacy and context dependency mean pretest scores were moderate. Attitude training should thus be regarded a vital and explicit starting point of inquiry-focused teacher professionalization (see also Van Aalderen-Smeets, Walma van der Molen, & Asma, 2012). Second, as shown, attitude change and school practice reform do not happen overnight and that, therefore, inquiry-focused school development requires extensive teacher professional development. Although schools likely prefer to participate in few,

short and hands-on workshops on inquiry teaching because of their limited time, attention and resources, our current findings suggest that such workshops will likely affect little significant change in what teachers think and do. And third, we underscore the importance of a 'school-wide' approach to the implementation of inquiry-based pedagogy in primary schools, as such an approach stimulates teacher collaboration and the cultivation of shared positive norms on practice reform. In sum, these findings thus provide empirical evidence to support the efficacy of the main guidelines formulated by the Exploratory Committee for the implementation of inquiry-based pedagogy in Dutch primary schools (see chapter 1).

Nevertheless, the present findings also give rise to several recommendations for improving our program. In the view of many teachers (study 3), the stimulation of pupils' higher-order thinking appeared to be one of the most difficult and challenging tasks of their professional development. In particular, teachers indicated to struggle with creatively 'responding' to pupils' inquisitive questions and ideas in the midst of lesson activities (as opposed to preparing open-ended, thematic study assignments or research projects for their pupils to work on). As suggested by Miri, David and Uri (2007), this competency may require that primary teachers possess sufficient higher-order thinking skills themselves – to be able to creatively and promptly respond to pupils' creativity and curiosity during lesson activities – which is a skill that cannot be easily taught. We thus recommend that the current program should be extended with follow-up coaching to help primary teachers master this complex inquiry teaching skill, on the job, over a longer period of time. Several teachers stated during our interviews to desire such coaching.

In addition, the findings of study 3 indicate that school principals' leadership development was not implicated through their participation in inquiry-focused teacher professionalization. This shortcoming was perceived by several school principals too, who therefore indicated to desire specialized leadership training on how to encourage and guide teachers' adoption of inquiry-based pedagogy. In line with research on school leadership, we recommend that such training should not include participation in stand-alone (university) leadership courses, but rather involve participation in formal mentoring programs in their own school (Grissom & Harrington, 2010). Under guidance of expert leaders, school principals may be encouraged and supported to set meaningful performance goals and to devise strategies at the school level to achieve these goals (see also Murphy & Seashore, 2018). For example, to re-allocate teachers' time to offer them the opportunity for

teacher collaboration and classroom experimentation or to role-model to teachers what it means to be 'inquiry-oriented' towards reform.

Lastly, our findings suggest that teachers' attitude development is dependent on the extent to which teachers collectively reflect on and discuss their learning experiences *during* training. This finding is in line with general recommendations for teacher professional development, which indicate that active and collective participation during professionalization is important (e.g., Desimone, 2009). We thus recommend that trainers and school principals regularly stimulate such reflection and discussion during course meetings. This can be accomplished by incorporating exercises and assessments that aim to raise participants' awareness of their own personal beliefs and feelings towards their practice reform, and challenge them to discuss these.

Opportunities to shorten inquiry-focused teacher professionalization

The above-described features of and recommendations for improving the present school improvement program may help advance the effectiveness of inquiry-focused school interventions. However, our current findings also suggest various opportunities to shorten and focus such professionalization. This would likely appeal to primary schools, as most primary teachers have limited time, attention and resources available for training and practice reform.

First, we found that primary teachers' creative lesson design perceptions were already quite positive at the beginning of the study. We note that teachers' positive creative lesson design scores may be the result of self-selection bias, as school teams that already held positive perceptions about this matter may have felt more inclined to enroll into the current program in the first place (to learn *how* to do so). This assumption should be tested. The present course could perhaps be shortened by spending less time on the positive development of teachers' creative lesson design perceptions.

Second, in line with Van Aalderen-Smeets and Walma van der Molen (2015), our findings indicate that primary teachers' self-efficacy and context dependency beliefs functioned as significant drivers of their inquiry teaching behavior. The current training course could thus perhaps be focused on the positive development of teachers' self-efficacy and context dependency beliefs alone (excluding a focus on teachers' perceived relevance, difficulty, enjoyment of inquiry teaching).

Third, the findings of the fourth study (chapter 5) indicate that pupils' attitudes towards curiosity and their implicit ability beliefs may function as particularly salient

drivers of their efforts and motivations to be inquiry-driven learners. Furthermore, as part of these attitudes and beliefs, pupils' perceptions about the value of curiosity for their own learning (Personal Inclination) and about the malleability of their intellectual ability (Incremental Beliefs) appeared to be most influential. Teachers' professionalization could thus perhaps be further focused on helping teachers stimulate and assess these attitude and motivation components among their pupils in particular.

However, future studies should look into the effectiveness of focused teacher enhancement, such as versions proposed above. It remains to be seen whether one or few components of teachers' and pupils' attitudes, beliefs and motivations can be fostered independently of other components. For example, the positive development of pupils' Personal Inclination perception might be dependent on the development of their Self-Efficacy perception (i.e., to be able to perceive the learning-value of curiosity, one has to first feel confident in being curious). The same interdependence may apply to other attitude and motivation components.

Alternatively, we propose that pre-service teacher training programs should emphasize the positive development of preservice teachers' attitudes towards inquiry teaching (and their creative lesson design perceptions). This would relieve in-service teacher training programs from spending time on achieving such attitude change, which would leave considerably time to focus on the development of teachers' didactical knowledge and skills for encouraging and assessing pupils' inquiry as part of day-to-day classroom practice (see also recommendations by De Vries, Van Keulen, Peters, & Walma van der Molen, 2011).

Concluding remarks

The main findings of the present study indicate that a school-wide approach to the implementation of S&T education, coupled with extensive attitude-focused and didactical professional training that adheres to a broader focus on S&T education, can prepare primary schools to integrate S&T education into their school organization. This dissertation thereby provides empirical evidence to support the efficacy of the main guidelines proposed by the Exploratory Committee (Verkenningscommissie, 2013) for the successful implementation of S&T teaching practice in Dutch primary education (see chapter 1).

At the same time, our findings highlight that it is not easy for primary teachers to implement inquiry-based pedagogy on their own. Changing habits is difficult,

especially when change requires the mastery of new complex teaching skills. In our view, scientific progress in this field can only be achieved when researchers are aware of the complexity of S&T-focused interventions. They should make explicit and substantiated decisions about what factors to foster and measure, at multiple school levels, and use evidence-based interventions and methodologically sound instruments wherever possible. Therefore, more longitudinal quantitative and qualitative research is needed to gain further insight into the impact of inquiry-focused school interventions and to further explore the processes of implementation. In particular, more attention should be given to the development of instruments to measure relevant changes in pupils', teachers', and school principals' knowledge, skills, attitudes and behaviors. These instruments may also serve as guides for what variables are essential to foster, and allow comparison of results across studies, to further build our knowledge base.

Other important challenges remain as well. For inquiry-focused school interventions to move beyond isolated and promising examples towards more widespread educational change, larger systemic issues and policies will need to be considered. For example, the incorporation of the positive development of pupils' images of and attitudes towards curiosity into national education standards (as advocated in this dissertation) would provide a strong incentive for school districts and teachers to emphasize its achievement as part of daily school practice (see also Pellegrino & Hilton, 2012).

To that end, in The Netherlands, a national consortium of teachers and school principals was recently charged with the task to redesign a national curriculum framework of what knowledge, skills, and attitudes Dutch pupils need to achieve by the end of primary and secondary education (In Dutch: the 'Curriculum.nu' project; see Onderwijsraad, 2014, 2016). This consortium was supported by the Dutch Department of Education. In the development of this (and other forthcoming) frameworks, we recommend the inclusion of education standards that call on the range of pupil qualities important for inquiry in all of the major content areas (not just applicable to the domains of science and technology), such as pupils' attitudes towards curiosity and their implicit inquiry ability beliefs. Foundations and federal officials should therefore promote research aimed at more clearly defining and developing assessments of such pupil qualities, initially for research purposes, and later for summative and formative assessment (see also Geisinger, 2016; Inspectie van het Onderwijs, 2019).

Only when pupils' inquiry is perceived by primary teachers as integral to school learning, no less so than solving a mathematical problem or writing an essay, will we likely see significant changes happen in classroom practice. Teachers need to become convinced about the use, pleasure, and opportunity of learning by inquiry themselves, which will empower them to act as genuine inquiry-driven role-models to their pupils. In turn, they should cultivate in pupils a sense that school is not only about memorizing historical facts, understanding how to gather information from the internet, or applying arithmetic to calculate surface areas. Rather, that school is ultimately about becoming an inquisitive explorer, an original thinker, and a resilient competitor in whatever future field they may find themselves. Teachers will have to explain to pupils why their teachers will keep challenging them with novelty and complexity throughout their academic careers, what they will have to put in to excel in life, and what they can expect to gain if they truly commit themselves.

This way, school can become a place where pupils' inquisitive, creative and bold thinking not only survives, but flourishes.

References

References

Abd-El-Khalick, F. (2012). Examining the sources for our understandings about science: Enduring conflations and critical issues in research on nature of science in science education. *International Journal of Science Education, 34*, 353–374.

Aikenhead, G. S. (2006). *Science education for everyday-life: Evidence-based practice*. New York: Teachers College Press.

Ainley, M. (2006). Connecting with learning: Motivation, affect and cognition in interest processes. *Educational Psychology Review*, *18*, 391–405.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50,* 179–211.

Ajzen, I. (2001). Nature and operation of attitudes. *Annual Review of Psychology, 52*(1), 27–58.

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood-Cliffs, NJ: Prentice-Hall.

Ajzen, I., & Fishbein, M. (2005). The influence of attitudes on behavior. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 173–221). Mahwah, NJ: Erlbaum.

Akerson, V. L., Abd-El-Khalick, F., & Lederman, N. G. (2000). Influence of a reflective explicit activity-based approach on elementary teachers' conceptions of nature of science. *Journal of Research in Science Teaching*, *37*, 295–317.

Amabile, T. M., & Pillemer, J. (2012). Perspectives on the social psychology of creativity. *The Journal of Creative Behavior*, *46*(1), 3–15.

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*, 261–271.

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives; abridged edition*. NY: Addison Wesley Longman, Inc.

Anderson, S., Leithwood, K., & Strauss, T. (2010). Leading data use in schools: Organizational conditions and practices at the school and district levels. *Leadership and Policy in Schools, 9*, 292–327.

Arango-Muñoz, S. (2014). The nature of epistemic feelings. *Philosophical Psychology*, *27*, 193–211.

Armitage, C. J., & Conner, M. (2001). Efficacy of the theory of planned behaviour: A meta-analytic review. *British Journal of Social Psychology, 40*, 471–499.

References

Avery, L. M., & Meyer, D. Z. (2012). Teaching science as science is practiced: Opportunities and limits for enhancing preservice elementary teachers' self-efficacy for science and science teaching. *School Science and Mathematics, 112*, 395–409.

Baehr, J. (2013). Educating for intellectual virtues: From theory to practice. *Journal of Philosophy of Education*, *47*, 248–262.

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

Baumeister, R. F. (2016). Towards a general theory of motivation: Problems, challenges, opportunities, and the big picture. *Motivation and Emotion*, *40*(1), 1–10.

Bennett, J., Lubben, F., & Hogarth, S. (2007). Bringing science to life: A synthesis of the research evidence on the effects of context-based and STS approaches to science teaching. *Science Education, 91*, 347–370.

Berliner, D. C. (2002). Educational research: The hardest science of all. *Educational Researcher, 31*, 18–20.

Berlyne, D. E. (1954). A theory of human curiosity. *British Journal of Psychology, 45*, 180–191.

Berlyne, D. E. (1960). *Conflict, Arousal, and Curiosity*. New York: McGraw-Hill.

Berlyne, D. E. (1978). Curiosity and learning. *Motivation and Emotion, 2*, 97–175.

Biggers, M., & Forbes, C. T. (2012). Balancing teacher and student roles in elementary classrooms: Preservice elementary teachers' learning about the inquiry continuum. *International Journal of Science Education*, *34*, 2205–2229.

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, *78*, 246–263.

Blalock, C. L., Lichtenstein, M. J., Owen, S., Pruski, L., Marshall, C., & Toepperwein, M. (2008). In pursuit of validity: A comprehensive review of science attitude instruments 1935–2005. *International Journal of Science Education, 30*, 961–977.

Block, M. E. (1995). Development and validation of the children's attitudes towards integrated physical education–revised (CAIPE–R) inventory. *Adapted Physical Activity Quarterly, 12*, 60–77.

Boone Jr, H. N., & Boone, D. A. (2012). Analyzing Likert data. *Journal of Extension, 50*(2), 1–5.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational researcher*, *33*, 3–15.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How People Learn.* Washington, DC: National Academy Press.

Bricker, L. A., & Bell, P. (2008). Conceptualizations of argumentation from science studies and the learning sciences and their implications for the practices of science education. *Science Education, 92*, 473–498.

Bruner, J. S. (1961). The art of discovery. *Harvard Educational Review, 31*, 21–32.

Burnette, J. L., O'Boyle, E. H., VanEpps, E. M., Pollack, J. M., & Finkel, E. J. (2013). Mind-sets matter: A meta-analytic review of implicit theories and self-regulation. *Psychological Bulletin, 139*, 655–701.

Byman, R. (2005). Curiosity and sensation seeking: A conceptual and empirical examination. *Personality and Individual Differences, 38*, 1365–1379.

Capps, D. K., Crawford, B. A., & Constas, M. A. (2012). A review of empirical literature on inquiry professional development: Alignment with best practices and a critique of the findings. *Journal of Science Teacher Education, 23*, 291–318

Carter, S. R. (2016). Using confirmatory factor analysis to manage discriminant validity issues in social pharmacy research. *International Journal of Clinical Pharmacy, 38*, 731–737.

Chambers, C. & Johnston, C. (2002). Developmental differences in children's use of rating scales. *Journal of Pediatric Psychology, 27*, 27–36.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233–255.

Chouinard, M. M. (2007). Children's questions: A mechanism for cognitive development. *Society for Research in Child Development, 27*, 1–126.

Claxton, G. (2007). Expanding young people's capacity to learn. *British Journal of Educational Studies, 55*, 115–134.

Claxton, G. & Carr, M. (2004). A framework for teaching learning: The dynamics of dispositions. *Early Years, 24*, 87–97.

Conner, M., & Armitage, C. J. (1998). Extending the theory of planned behavior: A review and avenues for further research. *Journal of applied social psychology, 28*, 1429–1464.

Cook, C., Goodman, N. D., & Schulz, L. E. (2011). When science starts: Spontaneous experiments in preschoolers' exploratory play. *Cognition, 3*, 341–349.

Cooke, R., & Sheeran, P. (2004). Moderation of cognition-intention and cognition-behaviour relations: A meta-analysis of properties of variables from the theory of planned behaviour. *British Journal of Social Psychology, 43*, 159–186.

Coulson, R. (1992). Development of an instrument for measuring attitudes of early childhood educators towards science. *Research in Science Education, 22*, 101–105.

Darnon, C., Harackiewicz, J. M., Butera, F., Mugny, G., & Quiamzade, A. (2007). Performance-approach and performance-avoidance goals: When uncertainty makes a difference. *Personality and Social Psychology Bulletin*, *33*, 813–827.

David, D. B. & Witryol, S. L. (1990). Gender as a moderator variable in the relationship between an intrinsic motivation scale and short-term novelty in children. *The Journal of Genetic Psychology, 151*, 153–167.

Davidson, S. K., Passmore, C., & Anderson, D. (2010). Learning on zoo field trips: The interaction of the agendas and practices of students, teachers, and zoo educators. *Science Education, 94*, 122–141.

De Castella, K., & Byrne, D. (2015). My intelligence may be more malleable than yours: the revised implicit theories of intelligence (self-theory) scale is a better predictor of achievement, motivation, and student disengagement. *European Journal of Psychology and Education, 30*, 245–267.

Deci, E. L. (1975). *Intrinsic Motivation*. New York: Plenum.

Dede, C. (2010). Comparing frameworks for 21st century skills. *21st century skills: Rethinking how students learn*, *20*, 51–76.

Deng, F., Chen, D. T., Tsai, C. C., & Chai, C. S. (2011). Students' views of the nature of science: A critical review of research. *Science Education*, *95*, 961-999.

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

Dewey, J. (1910). *How we think*. New York: Heath.

DeWitt, J., & Storksdieck, M. (2008). A short review of school field trips: Key findings from the past and implications for the future. *Visitor Studies, 11,* 181–197.

Dinger, F. C., Dickhäuser, O., Spinath, B., & Steinmayr, R. (2013). Antecedents and consequences of students' achievement goals: A mediation analysis. *Learning and Individual Differences, 28*, 90–101.

Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality, and development*. Philadelphia: Psychology Press.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological review*, *95*, 256–273.

Eagly, A., & Chaiken, S. (1993). *The psychology of attitudes*. Belmont, CA: Wadsworth group/Thomson Learning.

Eisinga, R., Te Grotenhuis, M., & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? *International Journal of Public Health, 58*, 637–642.

Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist, 34*, 169–189.

Elliot, A. J., & Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology, 72*, 218–232.

Elliot, A. J., & McGregor, H. A. (2001). A 2x2 achievement goal framework. *Journal of personality and social psychology, 80*, 501–519.

Engel, S. (2006). Open Pandora's Box: Curiosity and imagination in the classroom. *Sarah Lawrence Occasional Paper Series*, 1–10.

Engel, S. (2009). Is curiosity vanishing? *Journal of the American Academy of Child and Adolescent Psychiatry, 48*, 777–779.

Engel, S. (2011). Children's need to know: Curiosity in schools. *Harvard Educational Review, 81*, 625–645.

Engel, S. (2013). The Case for Curiosity. *Educational Leadership, 70*, 36–40.

Engel, S. (2015). *The hungry mind: The origins of curiosity in childhood*. Harvard University Press.

Engel, S. & Randall, K. (2009). How teachers respond to children's inquiry. *American Educational Research Journal, 46*, 183–202.

Fadel, C. (2008). *21st Century skills: How can you prepare students for the new global economy?* Paper presented at the OECD/CERI.

Fadlelmula, F. K. (2010). Educational motivation and students' achievement goal orientations. *Procedia Social and Behavioral Sciences, 2*, 859–863.

Fishbein, M., & Ajzen, I. (1974). Attitudes towards objects as predictors of single and multiple behavioral criteria. *Psychological Review, 81*, 59–74.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 7*, 286–299.

Ford, M. E. (1992). *Human motivation: Goals, emotions, and personal agency beliefs*. Newbury Park, CA: Sage

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*, 39–50.

References

Fortus, D. (2014). Attending to affect. *Journal of Research in Science Teaching, 51,* 821–835.

Fouad, K. E., Masters, H., & Akerson, V. L. (2015). Using history of science to teach nature of science to elementary students. *Science & Education, 24,* 1103–1140.

Frantom, C. G., Green, K. E., & Hoffman, E. R. (2002). Measure development: The children's attitudes towards technology scale (CATS). *Journal of Educational Computing Research, 26,* 249–263.

Galen, B. R., & Underwood, M. K. (1997). A developmental investigation of social aggression among children. *Developmental Psychology, 33,* 589–600.

Glasman, L. R., & Albarracín, D. (2006). Forming attitudes that predict future behavior: A meta-analysis of the attitude-behavior relation. *Psychological Bulletin, 132,* 778–822.

Geisinger, K. F. (2016). 21st century skills: What are they and how do we assess them? *Applied Measurement in Education, 29,* 245–249.

Godin, G., & Kok, G. (1996). The theory of planned behavior: a review of its applications to health-related behaviors. *American Journal of Health Promotion, 11,* 87–98.

Goe, L., Bell, C., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis.* Washington, DC: National Comprehensive Center for Teacher Quality.

Goodman, S., Jaffer, T., Keresztesi, M., Mamdani, F., Mokgatle, D., Musariri, M., ... & Schlechter, A. (2011). An investigation of the relationship between students' motivation and academic performance as mediated by effort. *South African Journal of Psychology, 41,* 373–385.

Grissom, J. A., & Harrington, J. R. (2010). Investing in administrator efficacy: An examination of professional development as a tool for enhancing principal effectiveness. *American Journal of Education, 116,* 583–612.

Grossnickle, E. M. (2014). Disentangling curiosity: Dimensionality, definitions, and distinctions from interest in educational contexts. *Educational Psychology Review, 28,* 23–60.

Gruber, M. J., Gelman, B. D., & Ranganath, C. (2014). States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron, 84,* 486–496.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis* (6th ed.). Upper Saddle River, NJ: Prentice Hall.

Hargreaves, A., & Fullan, M. (2012). *Professional capital: Transforming teaching in*

every school. Teachers College Press.

Hargreaves, A., Lieberman, A., Fullan, M., & Hopkins, D. (Eds.). (2009). *Second international handbook of educational change*. Dordrecht, The Netherlands: Springer.

Hassan, M. M., Bashir, S., & Mussel, P. (2015). Personality, learning, and the mediating role of epistemic curiosity: A case of continuing education in medical physicians. *Learning and Individual Differences, 42*, 83–89.

Hausenblas, H. A., Carron, A. V., & Mack, D. E. (1997). Application of the theories of reasoned action and planned behavior to exercise behavior: A meta-analysis. *Journal of Sport and Exercise Psychology, 19*(1), 36–51.

Heywood, D., Parker, J., & Jolley, N. (2012). Pre-service teachers' shifting perceptions of cross-curricular practice: The impact of school experience in mediating professional insight. *International Journal of Educational Research, 55*, 89–99.

Hirschfeld, G., & von Brachel, R. (2014). Multiple-group confirmatory factor analysis in R–A tutorial in measurement invariance with continuous and ordinal indicators. *Practical Assessment, Research & Evaluation, 19*(7), 1–12.

Hodson, D. (2014). Learning science, learning about science, doing science: Different goals demand different learning methods. *International Journal of Science Education, 36*, 2534–2553.

Hong, Y. Y., Chiu, C. Y., Dweck, C. S., Lin, D., & Wan, W. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology, 77*, 588–599.

Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55.

Hunzicker, J. (2011). Effective professional development for teachers: A checklist. *Professional development in education, 37*, 177–179.

Inspectie van het Onderwijs (2019). *De Staat van het Onderwijs. Onderwijsverslag 2017/2018.* Utrecht: Inspectie van het Onderwijs.

Isen, A. M. (2000). Some perspectives on positive affect and self-regulation. *Psychological Inquiry, 11*, 184–187.

Jarvis, T., & Pell, A. (2004). Primary teachers' changing attitudes and cognition during a two-year science inservice program and their effect on pupils. *International Journal of Science Education, 26*, 1787–1811.

Jepma, M., Verdonschot, R. G., Van Steenbergen, H., Rombouts, S. A., & Nieuwenhuis, S. (2012). Neural mechanisms underlying the induction and relief of perceptual curiosity. *Memory and Motivational/Emotional Processes, 6*, article 5.

Jerald, C. D. (2009). Defining a 21" century education. Retrieved from the Center for Public Education website: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.460.8011&rep=rep1&type=pdf

Jirout, J. & Klahr, D. (2012). Children's scientific curiosity: In search of an operational definition of an elusive concept. *Developmental Review, 32*, 125–160.

Jirout, J. J. (2011). *Curiosity and the development of question generation skills.* Paper presented at the 2011 AAAI 2011 Symposium. Retrieved March 23rd, 2017, from https://www.aaai.org/ocs/index.php/FSS/FSS11/paper/viewFile/4194/4490

Jones, M. G., Jones B. D., & Hargrove, T. (2003). *The unintended consequences of high-stakes testing.* Lanham, MD: Rowman & Littlefield.

Jones, M. T., & Eick, C. J. (2007). Implementing inquiry kit curriculum: Obstacles, adaptations, and practical knowledge development in two middle school science teachers. *Science Education, 91*, 492–513.

Jones, M., & Shelton, M. (2011). *Developing Your Portfolio–Enhancing Your Learning and Showing Your Stuff: A Guide for the Early Childhood Student Or Professional.* New York: Taylor & Francis.

Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T.-Y., & Camerer, C. F. (2009). The wick in the candle of learning: epistemic curiosity activates reward circuitry and enhances memory. *Psychological Science, 20*, 963–973.

Kashdan, T. B. (2004). Curiosity. In C. Peterson & M. E. P. Seligman (Eds.), *Character strengths and virtues: a handbook and classification* (pp. 125–141). New York: Oxford University Press.

Kashdan, T. B. & Roberts, J. E. (2004). Trait and state curiosity in the genesis of intimacy: Differentiation from related constructs. *Journal of Social and Clinical Psychology, 23*, 792–816.

Kashdan, T. B. & Steger, M. F. (2007). Curiosity and pathways to well-being and meaning in life: Traits, states, and everyday behaviors. *Motivation and Emotion, 31*, 159–173.

Kashdan, T. B., & Silvia, P. J. (2009). Curiosity and interest: The benefits of thriving on novelty and challenge. *Oxford handbook of positive psychology*, *2*, 367–374.

Kashdan, T. B., Gallagher, M. W., Silvia, P. J., Winterstein, B. P., Breen, W. E., Terhar, D., & Steger, M. F. (2009). The curiosity and exploration inventory-II: Development, factor structure, and psychometrics. *Journal of Research in Personality, 43*, 987–998.

Kashdan, T. B., Rose, P., & Fincham, F. D. (2004). Curiosity and exploration: facilitating positive subjective experiences and personal growth opportunities. *Journal of Personality Assessment, 82*, 291–305.

Kennedy, M. (2016). Parsing the practice of teaching. *Journal of Teacher Education*, *67*(1), 6-17.

Khishfe, R., & Abd-El-Khalick, F. (2002). Influence of explicit and reflective versus implicit inquiry-oriented instruction on sixth graders' views of nature of science. *Journal of Research in Science Teaching*, *39*, 551–578.

King, F., Goodson, L., & Rohani, F. (2011). *Higher order thinking skills: Definitions, strategies, assessment.* Center for Advancement of Learning and Assessment. Tallahassee, FL: Florida State University.

Krosnick J. A., & Fabrigar, L. R. (1997). Designing rating scales for effective measurement in surveys. In L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz, and D. Trewin (Eds.), *Survey measurement and process quality* (pp. 141–164). New York: Wiley.

Kulas, J. T., & Stachowski, A. A. (2009). Middle category endorsement in odd-numbered Likert response scales: Associated item characteristics, cognitive demands, and preferred meanings. *Journal of Research in Personality, 43*, 489–493.

Lederman, N. & Abell, S. K. (eds.) (2014). *Handbook of Research on Science Education.* Volume II: Routledge.

Lederman, N. G., Antink, A., & Bartos, S. (2014). Nature of science, scientific inquiry, and socio-scientific issues arising from genetics: A pathway to developing a scientifically literate citizenry. *Science & Education*, *23*, 285–302.

Leithwood, K., Harris, A., & Hopkins, D. (2008). Seven strong claims about successful school leadership. *School leadership and Management, 28*, 27–42.

Levinsen, K., & Nielsen, J. (2011). Innovating design for learning in the networked society. In L. Dirckinck-Holmfeld, V. Hodgson, & D. McConnell (Eds.), *Exploring the theory, pedagogy and practice of networked learning* (pp. 237-256). New York: Springer Publishing Company.

Litman, J. A. (2008). Interest and deprivation dimensions of epistemic curiosity. *Personality and Individual Differences, 44,* 1585–1595.

Litman, J. A. & Pezzo, M. V. (2007). Dimensionality of interpersonal curiosity. *Personality and Individual Differences, 43,* 1448–1459.

Litman, J. A. & Spielberger, C. D. (2003). Measuring epistemic curiosity and its diversive and specific components. *Journal of Personality Assessment, 80,* 75–86.

Litman, J. A., & Pezzo, M. V. (2007). Dimensionality of interpersonal curiosity. *Personality and Individual Differences, 43,* 1448–1459.

Litman, J. A., & Spielberger, C. D. (2003). Measuring epistemic curiosity and its diversive and specific components. *Journal of Personality Assessment, 80,* 75–86.

Litman, J. A., Crowson, H. M., & Kolinski, K. (2010). Validity of the interest-and deprivation-type epistemic curiosity distinction in non-students. *Personality and Individual Differences*, *49,* 531–536.

Litman, J. A., Robinson, O. C., & Demetre, J. D. (2016). Intrapersonal curiosity: Inquisitiveness about the inner self. *Self and Identity, 16,* 231–250.

Litman, J., Hutchins, T., & Russon, R. (2005). Epistemic curiosity, feeling-of-knowing, and exploratory behaviour. *Cognition & Emotion, 19,* 559–582.

Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83,* 1198–1202.

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin, 116*(1), 75–98.

Logsdon, L. F. (2013). Questioning the role of "21st-century skills" in arts education advocacy discourse. *Music Educators Journal*, *100,* 51–56.

Loucks-Horsley, S. (2001). The Concerns-Based Adoption Model (CBAM): A model for change in individuals. In: Bybee R (Ed.), *National Standards & the Science Curriculum.* Kendall/Hunt Publishing Co: Dubuque.

Lucas, B., Claxton, G., & Spencer, E. (2013). *Progression in student creativity in school: First steps towards new forms of formative assessments* (OECD Education Working Paper No. 86). Paris, France: OECD.

Luce, M. R., & Hsi, S. (2014). Science-relevant curiosity expression and interest in science: an exploratory study. *Science Education*, *99*(1), 70–97.

Lumpe, A. T., Czerniak, C. M., Haney, J. J., & Beltyukova, S. (2012). Beliefs about teaching science: The relationship between elementary teachers' participation

in professional development and student achievement. *International Journal of Science Education, 34,* 153–166.

Lumpe, A.T., Haney, J.J., & Czerniak, C.M. (2000). Assessing teachers' beliefs about their science teaching context. *Journal of Research in Science Teaching, 37,* 275–292.

Luo, W., Paris, S. G., Hogan, D., & Luo, Z. (2011). Do performance goals promote learning? A pattern analysis of Singapore students' achievement goals. *Contemporary Educational Psychology, 36,*165–176.

Makar, K., & Fielding-Wells, J. (2018). Shifting more than the goal posts: developing classroom norms of inquiry-based learning in mathematics. *Mathematics Education Research Journal, 30,* 53–63.

Mangels, J. A., Butterfield, B., Lamb, J., Good, C., & Dweck, C. S. (2006). Why do beliefs about intelligence influence learning success? A social cognitive neuroscience model. *Social Cognitive and Affective Neuroscience, 1,* 75–86.

Marx, R. W. & Harris, C. J. (2006). No Child Left Behind and science education: Opportunities, challenges, and risks. *The Elementary School Journal, 106,* 467–478.

Marzano, R., Pickering, D., & Pollock, J. (2001). *Classroom instruction that works.* Alexandria, VA: Association for Supervision and Curriculum Development.

Maw, W. H. & Maw, E. W. (1964). *An exploratory investigation into the measurement of curiosity in elementary school children.* Cooperative Research Project 801. University of Delaware.

McCombs, B. L., Daniels, D. H., & Perry, K. E. (2008). Children's and teachers' perceptions of learner-centered practices, and student motivation: Implications for early schooling. *The Elementary School Journal, 109,* 16–35.

Mellor, D., & Moore, K. A. (2014). The use of Likert scales with children. *Journal of Pediatric Psychology, 39,* 369–379.

Metz, K. E. (2008). Narrowing the gulf between the practices of science and the elementary school science classroom. *The Elementary School Journal, 109,* 138–161.

Milfont, T. L., & Fischer, R. (2015). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research, 3,* 111–130.

Miller, R., Goddard, R., Kim, M., Jacob, R., Goddard, Y. & Schroeder, P. (2016). Can Professional Development Improve School Leadership? Results From a Randomized Control Trial Assessing the Impact of McREL's Balanced

Leadership Program on Principals in Rural Michigan Schools. *Educational Administration Quarterly, 52*, 531–566.

Miri, B., David, B. C., & Uri, Z. (2007). Purposely teaching for the promotion of higher-order thinking skills: A case of critical thinking. *Research in Science Education, 37*, 353–369.

Moolenaar, N. M., Sleegers, P. J. C., & Daly, A. J. (2012). Teaming up: Linking collaboration networks, collective efficacy, and student achievement. *Teaching and Teacher Education, 28*, 251–262.

Mueller, R. O., & Hancock, G. R. (2008). Best practices in structural equation modeling. In J. W. Osborne (Ed.), *Best practices in quantitative methods* (pp. 488–508). Thousand Oaks, CA: Sage.

Muijs, D. (2011). Leadership and organisational performance: From research to prescription? *International Journal of Educational Management, 25*(1), 45–60.

Muis, K. R., Pekrun, R., Sinatra, G. M., Azevedo, R., Trevors, G., Meier, E., & Heddy, B. C. (2015). The curious case of climate change: Testing a theoretical model of epistemic beliefs, epistemic emotions, and complex learning. *Learning and Instruction, 39*, 168–183.

Muis, K. R., Psaradellis, C., Lajoie, S. P., Di Leo, I., & Chevrier, M. (2015). The role of epistemic emotions in mathematics problem solving. *Contemporary Educational Psychology, 42*, 172–185.

Murphy, J. & Seashore Louis, K (2018). *Positive school leadership: Building capacity and strengthening relationships*. New York, NY: Teachers College Press.

Mussel, P. (2010). Epistemic curiosity and related constructs: Lacking evidence of discriminant validity. *Personality and Individual Differences, 49*, 506–510.

Muthén, L.K. and Muthén, B.O. (1998-2015). *Mplus User's Guide*. Seventh Edition. Los Angeles, CA: Muthén & Muthén.

National Research Council (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Committee on a conceptual framework for new K-12 Science Education Standards. Board on science education Division of behavioral and social sciences and education. Washington, DC: The National Academies Press.

Next Generation Science Standards (2013). *Next Generation Science Standards: For states, by states*. Retrieved from http://www.nextgenscience.org.

Olson, J. M., & Zanna, M. P. (1993). Attitudes and attitude change. *Annual review of psychology, 44*, 117–154.

Olson, S. & Loucks-Horsley, S. (2000). *Inquiry and the National Science Education Standards: A Guide for Teaching and Learning.* Washington, DC: National Academy Press.

Onderwijsraad (2014). *Een eigentijds curriculum.* Den Haag: Onderwijsraad.

Onderwijsraad (2016). *Over de volle breedte van onderwijskwaliteit. Van smal beoordelen naar breed verantwoorden.* Den Haag: Onderwijsraad.

Opdal, P. M. (2001). Curiosity, wonder and education seen as perspective development. *Studies in Philosophy and Education, 20*, 331–344.

Organization For Economic Co-Operation And Development [OECD] (2015). *OECD skills outlook 2015: Youth, skills and employability.* Paris, France: OECD.

Osborne, J. (2014). Teaching scientific practices: Meeting the challenge of change. *Journal of Science Teacher Education, 25*, 177–196.

Osborne, J., & Dillon, J. (2008). *Science education in Europe: Critical reflections (Vol. 13).* London: The Nuffield Foundation.

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education, 25,* 1049–1079.

Palmer, D. H. (2006). Sources of self-efficacy in a science methods course for primary teacher education students. *Research in Science Education, 36*, 337–353.

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas.* New York: Basic Books.

Paunesku, D., Walton, G. M., Romero, C., Smith, E. N., Yeager, D. S., & Dweck, C. S. (2015). Mind-set interventions are a scalable treatment for academic underachievement. *Psychological Science, 26*, 784–793.

Pekrun, R. (2006). The control-value theory of achievement emotions: assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*, 315–341.

Pekrun, R., Goetz, T., Frenzel, A., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: the Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology, 36*, 36–48.

Pellegrino J. & Hilton M.L. (2012). *Education for Life and Work. Transferable Knowledge and Skills for the 21st Century.* Washington, D.C.

Piaget, J. (1952). *The origins of intelligence in children.* New York: International Universities Press.

Piaget, J. (1969). *The Psychology of Intelligence.* New York: Littlefield, Adams.

Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of educational Psychology, 95*, 667–686.

Piotrowski, J. T., Litman, J. A., & Valkenburg, P. (2014). Measuring epistemic curiosity in young children. *Infant and Child Development, 23*, 542–553.

Platform Onderwijs2032 (2016). *Ons onderwijs2032: Eindadvies.* Den Haag: Bureau Platform Onderwijs2032.

Pluck, G. & Johnson, H. L. (2011). Stimulating curiosity to enhance learning. *GESJ: Education Sciences and Psychology, 2*(19), 24–31.

Post, T. & Walma van der Molen, J. H. (2014). Effects of company visits on Dutch primary school children's attitudes toward technical professions. *International Journal of Technology and Design Education, 24*, 349–373.

Post, T., & Walma van der Molen, J. H. (2018a). Do children experience curiosity at school? Exploring children's own conceptions of curiosity inside and outside the school context. *Learning, Culture and Social Interaction, 18,* 60–71.

Post, T., & Walma van der Molen, J. H. (2018b). Development and Validation of a Questionnaire to Measure Primary School Children's Images of and Attitudes Towards Curiosity (the CIAC Questionnaire). *Motivation and Emotion,* 1–20.

Post, T., & Walma van der Molen, J. H. (in press). Effects of a longitudinal school development program on primary teachers' attitudes towards inquiry teaching and their inquiry teaching practices. *Teacher College Records.*

Prudon, P. (2015). Confirmatory factor analysis as a tool in research using questionnaires: A critique. *Comprehensive Psychology, 4*, 10.

Ramey-Gassert, L., Shroyer, M. G., & Staver, J. R. (1996). A qualitative study of factors influencing science teaching self-efficacy of elementary level teachers. *Science Education, 80*, 283–315.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Thousand Oaks, CA: Sage Publications.

Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement, 21*, 173–184.

Reeve, J. (1989). The interest-enjoyment distinction in intrinsic motivation. *Motivation and Emotion, 13*, 83–103.

Reeve, J. (2015). *Understanding motivation and emotion* (6th ed.). Hoboken, NJ: John Wiley & Sons.

Reid, N. (2006). Thoughts on attitude measurement. *Research in Science & Technological Education, 24*(1), 3–27.

Reio Jr, T. G., Petrosko, J. M., Wiswell, A. K., & Thongsukmag, J. (2006). The measurement and conceptualization of curiosity. *The Journal of Genetic Psychology, 167*, 117–135.

Reise, S. P., Waller, N. G., & Comrey, A. L. (2000). *Factor analysis and scale revision. Psychological Assessment, 12*, 287–297.

Ricketts, A. (2014). Preservice elementary teachers' ideas about scientific practices. *Science & Education*, *23*, 2119–2135.

Rojas-Drummond, S., Maine, F., Alarcón, M., Trigo, A. L., Barrera, M. J., Mazón, N., Vélez, M., & Hofmann, R. (2017). Dialogic literacy: Talking, reading and writing among primary school children. *Learning, Culture and Social Interaction*, *12*, 45–62.

Rosnow, R. L. (2001). Rumor and gossip in interpersonal interaction and beyond: A social exchange perspective. In R. M. Kowalski (Ed.), *Behaving badly: Aversive behaviors in interpersonal relationships*. Washington, DC: American Psychological Association.

Sawyer, K. (2015). A call to action: The challenges of creative teaching and learning. *Teachers College Record, 117*(10), 1–34.

Sawyer, R. K. (2011). How to transform schools to foster creativity. *Teachers College Record, 118*(4), 1–23.

Schmitt, T. A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment, 29*, 304–321.

Schoon, K. J. & Boone, W. J. (1998). Self-efficacy and alternative conceptions of science of preservice elementary teachers. *Science Education, 82*, 553–568.

Schwarz, N. (2008). Attitude measurement. In W. Crano & R. Prislin (Eds.), *Attitudes and attitude change* (pp. 41–60). Philadelphia, PA: Psychology Press.

Senko, C., Hulleman, C. S., & Harackiewicz, J. M. (2011). Achievement goal theory at the crossroads: Old controversies, current challenges, and new directions. *Educational Psychologist, 46*, 26–47.

Shore, B. M., Aulls, M.W., & Delcourt, M. A. B. (Eds.) (2008). *Inquiry in education (vol. II): Overcoming barriers to successful implementation*. New York: Erlbaum.

Silvia, P. J. (2006). *Exploring the Psychology of Interest*. New York: Oxford University Press.

References

Silvia, P. J., & Sanders, C. E. (2010). Why are smart people curious? Fluid intelligence, openness to experience, and interest. *Learning and Individual Differences, 20*, 242–245.

Simon, H. A. (2001). "Seek and ye shall find": How curiosity engenders discovery. In K. Crowley, C. Schunn, & T. Okada (Eds.), *Designing for science: Implications from everyday, classroom, and professional settings* (pp. 5–20). Mahwah, NJ: Erlbaum.

Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. *Educational Psychologists, 50*, 1–13.

Slavin, R. E. (2007). *Educational research: In an age of accountability*. Boston: Pearson Education.

Slavin, R. E., Lake, C., Hanley, P., & Thurston, A. (2014). Experimental evaluations of elementary science programs: A best-evidence synthesis. *Journal of Research in Science Teaching, 51*, 870–901.

Southerland, S. A., Johnston, A., & Sowell, S. (2006). Describing teachers' conceptual ecologies for the nature of science. *Science Education*, *90*, 874–906.

Spektor-Levy, O., Baruch, Y. K., & Mevarech, Z. (2013). Science and Scientific Curiosity in Pre-school—The teacher's point of view. *International Journal of Science Education, 35*, 2226–2253.

Spencer, E., Lucas, B., & Claxton, G. (2014). Progression in student creativity in school: first steps towards new forms of formative assessments. *Contemporary Readings in Law and Social Justice, 6*, 81–121.

Spielberger, C. D., & Starr, L. M. (1994). Curiosity and exploratory behavior. In H. F. O'Neil & M. Drillings (Eds.), *Motivation: Theory and research* (pp. 221–243). Hillsdale, NJ: Erlbaum.

Stender, A., Schwichow, M., Zimmerman, C., & Härtig, H. (2018). Making inquiry-based science learning visible: the influence of CVS and cognitive skills on content knowledge learning in guided inquiry. *International Journal of Science Education*, *40*, 1812–1831.

Stoddart, T., Pinal, A., Latzke, M., & Canaday, D. (2002). Integrating inquiry science and language development for English language learners. *Journal of research in science teaching, 39*, 664–687.

Stuckey, M., Hofstein, A., Mamlok-Naaman, R., & Eilks, I. (2013). The meaning of 'relevance' in science education and its implications for the science curriculum. *Studies in Science Education, 49*(1), 1–34.

Syer, C. A., Chichekian, T., Shore, B. M., & Aulls, M. W. (2012). Learning "to do" and learning "about" inquiry at the same time: Different outcomes in valuing the importance of various intellectual tasks in planning, enacting, and evaluating an inquiry curriculum. *Instructional Science, 41*, 521–537.

Tai, R. H, Liu, C. Q, Maltese, A. V, & Fan, X. (2006). Planning early for careers in science. *Science, 312*, 1143–1145.

Tamdogon, O. G. (2006). Creativity in education: Clearness in perception, vigorousness in curiosity. *Education for information, 24*, 139–151.

Tempelaar, D. T., Rienties, B., Giesbers, B. & Gijselaers, W. H. (2015). The pivotal role of effort beliefs in mediating implicit theories of intelligence and achievement goals and academic motivations. *Social Psychology of Education, 18*, 101–120.

Teo, P. (2013). 'Stretch your answers': Opening the dialogic space in teaching and learning. *Learning, Culture and Social Interaction, 2*, 91–101.

Thijs, A., Fisser, P, & Van der Hoeven, M. (2014). *21^e eeuwse vaardigheden in het curriculum van het funderend onderwijs*. Enschede, The Netherlands: SLO Nationaal Expertisecentrum Leerplanontwikkeling.

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

Thurlings, M., Evers, A. T., & Vermeulen, M. (2015). Toward a model of explaining teachers' innovative behavior: A literature review. *Review of Educational Research, 85*, 430–471.

Timperley, H. (2008). *Teacher Professional Learning and Development*. Educational Practices Series, 18th edition. Brussels: International Academy of Education.

Trevors, G. J., Muis, K. R., Pekrun, R., Sinatra, G. M., & Muijselaar, M. M. (2017). Exploring the relations between epistemic beliefs, emotions, and learning from texts. *Contemporary Educational Psychology, 48*, 116–132.

Trochim, W. M., & Donnelly, J. P. (2006). *The research methods knowledge base* (3rd ed.). Cincinnati, OH: Atomic Dog.

Tsai, C. C. (2006). Reinterpreting and reconstructing science: Teachers' view changes toward the nature of science by courses of science education. *Teaching and Teacher Education*, *22*, 363–375.

Turner, S., & Ireson, G. (2010). Fifteen pupils' positive approach to primary school science: When does it decline? *Educational Studies, 36*, 119–141.

Ucar, S., & Demircioglu, T. (2011). Changes in preservice teacher attitudes toward astronomy within a semester-long astronomy instruction and four-year-long teacher training programme. *Journal of Science Education and Technology, 20*, 65–73.

Urick, A., & Bowers, A. J. (2014). What are the different types of principals across the United States? A latent class analysis of principal perception of leadership. *Educational Administration Quarterly, 50*, 96-134.

Uiterwijk-Luijk, L., Krüger, M., & Volman, M. (2019). Promoting inquiry-based working: Exploring the interplay between school boards, school leaders and teachers. *Educational Management Administration & Leadership*, *47*, 475–497.

Van Aalderen-Smeets, S.I., Walma van der Molen, J. H., & Asma, L. J. F. (2012). Primary teachers' attitudes towards science and technology: A new theoretical framework. *Science Education, 96*, 158–182.

Van Aalderen-Smeets, S., & Walma van der Molen, J. (2013). Measuring primary teachers' attitudes toward teaching science: Development of the dimensions of attitude toward science (DAS) instrument. *International Journal of Science Education, 35*, 577–600.

Van Aalderen-Smeets, S. I., & Walma van der Molen, J. H. (2015). Improving primary teachers' attitudes towards science by attitude-focused professional development. *Journal of Research in Science Teaching, 52*, 710–734.

Van Aalderen-Smeets, S. I., & Walma van der Molen, J. H. (2018a). Modeling the relation between students' implicit beliefs about their abilities and their educational STEM choices. *International Journal of Technology and Design Education, 28*(1), 1–27.

Van Aalderen-Smeets, S. I., Walma van der Molen, J. H., & Xenidou-Dervou, I. (2018b). Implicit STEM ability beliefs predict secondary school students' STEM self-efficacy beliefs and their intention to opt for a STEM field career. *Journal of Research in Science Teaching, 56*, 465–485.

Van Aalderen-Smeets, S. I., Walma van der Molen, J. H., van Hest, E. G. C., & Poortman, C. (2017). Primary teachers conducting inquiry projects: effects on attitudes towards teaching science and conducting inquiry. *International journal of science education*, *39*, 238–256.

Van Booven, C. D. (2015). Revisiting the Authoritative–Dialogic Tension in Inquiry-Based Elementary Science Teacher Questioning. *International Journal of Science Education, 37*, 1182–1201.

Van der Scheer, E. A., Bijlsma, H. J., & Glas, C. A. (2018). Validity and reliability of student perceptions of teaching quality in primary education. *School Effectiveness and School Improvement*, 1-21.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(1), 4–70.

Verkenningscommissie Wetenschap en Technologie Primair Onderwijs (2013). *Advies verkenningscommissie wetenschap en technologie primair onderwijs*. PO-Raad en Platform Bèta Techniek.

Vogel, T. & Wänke, M. (2016). *Attitudes and attitude change* (2nd ed.). London: Routledge.

Von Stumm, S., Hell, B., & Chamorro-Premuzic, T. (2011). The hungry mind intellectual curiosity is the third pillar of academic performance. *Perspectives on Psychological Science, 6*, 574–588.

Voogt, J. & Pareja Roblin, N. (2010). *21st Century skills*. Discussienota. Enschede: Universiteit Twente.

De Vries, M. J., Van Keulen, H., Peters, S., & Walma van der Molen, J. (2011). *Professional development for primary teachers in science and technology: The Dutch VTB-Pro project in an international perspective.* (International Technology Education Studies; Vol. 9). Rotterdam/Boston/Taipei: Sense Publishers.

Walan, S., Mc Ewen, B., & Gericke, N. (2016). Enhancing primary science: an exploration of teachers' own ideas of solutions to challenges in inquiry-and context-based teaching. *Education 3-13, 44*(1), 81–92.

Warren, B., Ballenger, C., Ogonowski, M., Rosebery, A. S., & Hudicourt-Barnes, J. (2001). Rethinking diversity in learning science: The logic of everyday sense-making. *Journal of Research in Science Teaching*, *38*(5), 529–552.

Wilson, S. M. (2013). Professional development for science teachers. *Science*, 340, 310–313

Woo, S. E., Harms, P. D., & Kuncel, N. R. (2007). Integrating personality and intelligence: Typical intellectual engagement and need for cognition. *Personality and Individual differences, 43*, 1635–1639.

World Economic Forum (2018). The Global Competitiveness Report 2018. Retrieved from: http://www3.weforum.org/docs/GCR2018/05FullReport/TheGlobalCompetitivenessReport2018.pdf

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist, 34*, 806–838.

Yu, C.-Y. (2002). *Evaluation of model fit indices for latent variable models with categorical and continuous outcomes* (Doctoral dissertation, University of California). Retrieved May 5th, 2017, from Mplus website http://www.statmodel.com/download/Yudissertation.pdf

Appendices

Appendices

Appendix A: Attitude questionnaire instrument

Table A

Items for measuring teachers' perceptions of control

Subscale	Item
Self-Efficacy	I am well able to deal with questions from pupils during their inquiry activities
	I have enough knowledge of inquiry to sufficiently implement this approach to teaching in class
	I have a sufficient command of inquiry to be able to support pupils well in investigating and designing in class
	If pupils do not reach a solution during inquiry-oriented assignments, I think I can succeed in helping them make further progress
Context Dependency	For me, the availability of a ready-to-use existing package of materials is essential to teaching inquiry in class
	For me, the availability of an inquiry teaching method is decisive for whether or not I will teach inquiry in class
	For me, the support of my colleagues and the school is decisive for whether or not I will teach inquiry in class

Response options to the items were 1–5, with 1 labelled 'totally disagree' and 5 labelled 'totally agree'.

Table B

Items for measuring teachers' perceptions of lesson design

Subscale	Item
Creative Lesson Design	I like teaching about content that requires me to think hard and allows me to figure things out
	I often ask myself in-depth questions to make the subject matter I teach more interesting for me
	I like to come up with assignments about topics that I would also like to learn more about myself
	I like teaching about challenging content because that is often more interesting for me
	I like to come up with inquiry-oriented assignments for pupils myself because I find it challenging to do so
	I like to come up with my own subject matter to teach about because I enjoy learning something new
Prescribed Lesson Design	I like to come up with easy subject matter to teach that I am sure I can manage
	I like teaching subject matter that has single right answers that are easy for me to memorize
	I would rather stick to teaching subject matter that is easy for me to accomplish

Response options to the items were 1–5, with 1 labelled 'totally disagree' and 5 labelled 'totally agree'

Appendix B: Inquiry teaching practice interview format

Table A

Items for measuring teachers' use of inquiry teaching

Subscale	Item (Teacher)	Item (Pupils)
Inquiry Teaching Frequency[a]	How often do you teach through inquiry on average?	How often do you work on inquiry during lessons on average?
Inquiry Teaching Duration[b]	How much time do you spend on teaching through inquiry on average, when doing so?	How much time do you spend on inquiry lesson activities on average?
Cross-Curricular Inquiry Focus[c]	Which subject domains are generally involved when doing inquiry projects?	Which subject domains are generally involved with inquiry projects?

[a] Response options to the items were 1–5, with 1 labelled 'yearly', 2 'quarterly', 3 'monthly', 4 'weekly' and 5 'daily'.
[b] Response options to the items were 1–5, with 1 labelled 'hour', 2 'few hours', 3 'half a day', 4 'one day' and 5 'multiple days'.
[c] Response options to the items were multiple choice, with (a) labelled 'science', (b) 'geography', (c) 'history', (d) 'social studies' and (e) 'mathematics and literacy'. The number of checked subject domains were summed up for each respondent to create a score of 1 – 5.

Table B

Items for measuring teachers' inquiry teaching quality

Subscale	Item (Teacher)	Item (Pupils)
Stimulation of Pupils' Attitudes towards Curiosity	I often stimulate pupils' attitudes towards curiosity by appreciating their curiosity-driven questions about the subject matter	How often does your teacher say that he/she appreciates your curiosity-driven questions about the subject matter?
	I often stimulate pupils' attitudes towards curiosity with assignments that make them come up with their own curiosity-driven questions	How often does your teacher provide assignments that require you to come up with your own curiosity-driven questions?
	I often stimulate pupils' attitudes towards curiosity by expressing my own curiosity with them during lesson activities	How often does your teacher show his/her own curiosity to learn, during lesson activities?
Stimulation of Pupils' Growth Mindset	I stimulate my pupils' growth mindset by praising their efforts when studying difficult lesson content	How often does your teacher praise someone for doing his or her best to complete a difficult assignment?
	I stimulate my pupils' growth mindset by tempting them to tackle challenging assignments	How often does your teacher challenge you to tackle more difficult assignments?
	I stimulate my pupils' growth mindset by appreciating their mistakes as a positive aspect of their learning	How often does your teacher say that making mistakes is not bad, but important to learn from?
Stimulation of Pupils' Higher-Order Thinking	I stimulate pupils' thinking with assignments that allow for multiple correct answers or solutions	How often do you work on assignments with several possible correct answers or solutions?
	I stimulate pupils' thinking by encouraging them to come up with their own original solutions or ideas	How often does your teacher ask you to come up with your own original solutions or ideas?
	I stimulate pupils' thinking with assignments that require them to determine the quality of their own work	How often does your teacher require you to determine for yourself whether the work you have done is good enough?

Table B continued

Inquiry Practice Evaluation	I evaluate my inquiry teaching practice to determine whether I did enough to stimulate pupils' attitudes towards curiosity	n.a.
	I evaluate my inquiry teaching practice to determine whether I did enough to stimulate pupils' growth mindset	
	I evaluate my inquiry teaching practice to determine whether I did enough to stimulate pupils' higher-order thinking	

Appendix C: School leadership interview format

Table A

Items for measuring school principals' leadership behavior

Subscale	Item (Teacher)	Item (School principal)
School Improvement Goals	I think that my school principal is currently very clear about the school policy's focus on inquiry teaching I think that my school principal is currently very clear about what I am expected to do to implement inquiry teaching practice I think that my school principal is currently very concrete about the next step we need to take to implement inquiry teaching practice further in the school	I think that I am very clear to the teacher team about the school policy's focus on inquiry teaching practice I think that I am very clear to the teacher team about what they are expected to do to implement inquiry teaching in their practice I think that I am very concrete about the next step to take to implement inquiry teaching practice further in the school
School Improvement Evaluation	I think my school principal is fully able to assess the quality of my inquiry teaching I think my school principal is fully involved with my everyday experiences of implementing inquiry teaching I think my school principal is fully informed about my needs to implement inquiry teaching practice further	I think I am fully able to assess the quality of teachers' inquiry teaching I think I am fully involved with teachers' everyday experiences of implementing inquiry teaching I think I am fully informed of teachers' needs to implement inquiry teaching practice further

Table A continued

Inquiry Practice Support	I think that my school principal currently does more than enough to guide me with the further implementation of inquiry teaching practice	I think that I currently do more than enough to guide the teacher team with the further implementation of inquiry teaching practice
	I think that my school principal currently does more than enough to provide collegial support for the further implementation of inquiry teaching	I think that I currently do more than enough to provide the teacher team with collegial support for the further implementation of inquiry teaching
	I think that my school principal currently does more than enough to provide time and space to implement inquiry teaching practice further	I think that I currently do more than enough to provide the teacher team with time and space to implement inquiry teaching practice further

Response options to the items were 1–5, with 1 labelled 'totally disagree' and 5 labelled 'totally agree'

Appendix D: Overview of the course content

Table A

Overview of the main content of each course meeting

Meeting	Main content
First part of the course	
Meeting 1 *October*	Participants' were made aware of their own views on inquiry and science through group discussions led by the teacher trainer Participants were taught about the relevance of having a positive attitude towards inquiry for potentially living a more successful, healthy and enjoyable life Participants were taught about the epistemic importance of 'inquiry' for knowledge development and innovation in general Participants were assigned a take-home assignment to keep a diary on their personal wonderments and curiosities about everyday life objects and natural phenomena
Meeting 2 *November*	Participants were taught how to formulate research questions based on their own wonderments and curiosities about everyday life objects and natural phenomena Participants were taught how to encourage and enhance pupils' inquiry during typical lesson activities with guiding questions that stimulate pupils' higher-order thinking Participants were taught about the relevance of inquiry-based (science) teaching for enhancing pupils' learning performance and their inquiry skills
Meeting 3 *December*	Participants were taught about methods of inquiry, such as conducting design experiments Participants were asked to conduct a predefined small-scale experiment during this meeting (figuring out how to blow the biggest bubbles of bubble-gum), to formulate hypotheses, set up corresponding experiments, and evaluate results Participants were made aware of the joy of 'doing inquiry' and stimulated to feel less anxious about inquiry Participants were assigned a take-home assignment to conduct a small-scale experiment with their own pupils
Meeting 4 *January*	Participants were familiarized with the higher-order thinking skills defined by the taxonomy of Bloom (i.e., analysing, evaluating and creating) Participants were taught about the relevance of stimulating pupils' higher-order thinking skills for inquiry Participants practiced several higher-order thinking skills themselves and were taught how to stimulate these skills in their pupils
Meeting 5 *February*	Participants were asked to reflect on all previous meetings and were to describe the changes and development of their own attitudes towards teaching (science) through inquiry Participants were asked to discuss their ideas and plans for teaching one inquiry-focused lesson in preparation of the next course meeting

Table A continued

Second part of the course

Meeting 6 *March*	Participants presented visual reports of their inquiry lesson to each other. The teacher trainer provided all participants feedback for improvement The school principal was encouraged to lead group discussions with the team about plans to further expand and integrate inquiry-based teaching into daily school practice
Meeting 7 *April*	The school principal started the meeting with a short presentation for the teacher team about the vision, goals, and expectations for adopting inquiry-based pedagogy in the school Participants were taught about the relevance of pupils' attitudes towards epistemic curiosity, higher-order thinking skills, and growth mindset for learning throughout the curriculum Participants were taught how to incorporate inquiry-focused lesson objectives and assessment criteria alongside their usual lesson objectives and assessment criteria Participants were taught strategies to foster pupils' attitudes towards epistemic curiosity, higher-order thinking skills, and growth mindset during regular lesson activities Participants were assigned a take-home assignment to teach pupils about the relevance of their attitudes towards epistemic curiosity, higher-order thinking skills, and growth mindsets
Meeting 8 *May*	Participants practiced to incorporate inquiry-focused lesson objectives and assessment criteria into their existing lesson plans (across the school curriculum) Participants practiced how to enhance or revise their existing lesson plans with opportunities for pupils to conduct inquiry and for themselves to guide and foster pupils' inquiry Participants were assigned a take-home assignment to implement one (practiced) lesson plan enhanced with inquiry-based pedagogy in their own classrooms
Meeting 9 *June*	Participants presented their implemented 'inquiry-infused' lesson plan to the entire teacher team. The teacher trainer provided all participants with feedback for improvement The school principal was asked to lead discussions among their team members about plans to further expand and integrate inquiry-based teaching into daily school practice

Appendix E: Survey instrument

Table A

Items for measuring attitude towards epistemic curiosity

Subscale	Code	Item
Personal Inclination	PI1	I really like to wonder about all the things I learn at school
	PI2	It is very important to me to come up with interesting questions at school, because then I learn more about the things around me
	PI3	It is very important to me to wonder about lots of things in class, because then I learn more about all sorts of different things.
	PI4	I really like to ask questions about all sorts of things in class
Societal Relevance	SR1	I think people who want to know a lot are very important to the economy of The Netherlands
	SR2	I think people who ask good questions have a big impact on society
	SR3	I think people who often come up with interesting questions are very important to society
Negative Opinion	NO1	I feel classmates are being stubborn when they always want to know all about everything in class
	NO2	I find classmates to be annoying when they ask a lot of smart questions in class
	NO3	I feel people who ask a lot of questions come across as dumb
Fear of Classmates' Negative Judgment	FJ1	I'm afraid that my classmates will think I'm a nerd if I ask a lot of smart questions in class
	FJ2	I'm afraid that my classmates will think it's stupid if I want to know more about something we're learning in class
	FJ3	I find it scary to show that I'd like to know more about a topic in class
Self-Efficacy	SE1	I am really good at coming up with smart questions in class
	SE2	I am really good at coming up with new questions about all sorts of topics in lessons at school
	SE3	I think I am really good at figuring out new things at school
	SE4	I am really good at coming up with smart questions about all sorts of subjects at school

Table B

Items for measuring implicit beliefs about the malleability of intelligence

Component	Code	Item
Incremental Beliefs	IB1	I believe that I can improve my ability to think
	IB2	I believe that I can gradually improve my ability to think
	IB3	I believe that I can improve my ability to think by working on increasingly more difficult assignments
	IB4	I believe that I can improve my ability to think by giving it my best effort
	IB5	I believe that I can always somewhat improve my ability to think
Entity Beliefs	EB1	I believe that I cannot change my ability to think, because that is how I was born
	EB2	I believe that I will always stay just as good at thinking, because I cannot change that
	EB3	I believe that I will always stay just as good at thinking, because that is set in my brain
	EB4	I believe that my ability to think is set and nothing can change it
	EB5	I believe that my ability to think is set in my brain, nothing can change it

Table C

Items for measuring effort beliefs

Component	Code	Item
Positive Effort Beliefs	PE1	I work harder on assignments that I find difficult, because that makes me learn the most
	PE2	I do my best when working on difficult homework assignments, because that makes me learn the most
	PE3	I give it my best effort when working on a harder assignment, because that helps me get it done better
	PE4	I really like to pick assignments that are somewhat more difficult, because this makes me learn something new

Table D

Items for measuring achievement goal orientation motivation

Subscale	Code	Item
Mastery Goal Orientation	MA1	I really like to learn something new about a school subject, because I find that to be interesting for myself
	MA2	I really like to learn something new about a school subject, because then I understand things better
	MA3	I find it really fun to understand a school subject as well as possible, because I like to increase my knowledge
	MA4	I find it really fun to understand a school subject, because that will help me later
	MA5	I really like to learn something new about a school subject, because I think it is important to understand things better
Performance-Avoidance Goal Orientation	PA1 PA2	I prefer to work on easier assignments, so I cannot get anything wrong
		I prefer to let others do the more difficult parts of an
	PA3 PA4	assignment, so nobody can see that I might not know what to do
		I prefer working on easier assignments, so I do not get bad grades
	PA5	I prefer learning about easier things, so my teacher does not see me make mistakes
		I prefer working on easier assignments, so I do not get low grades

Summary

English summary

Introduction

The dissertation describes the results of a two-year school improvement program in which the complete school staffs of six Dutch primary schools were trained to integrate inquiry-based pedagogy into daily school practice. International education policy documents increasingly promote primary science and technology (S&T) education. Scientific and technological innovations take place in a rapidly increasing rate and lead to the ongoing transformation of labor markets and societal structures (World Economic Forum, 2018). All young people will therefore have to become sufficiently familiar with S&T to be able to fully participate in society (National Research Council, 2012). Furthermore, research indicates that children's natural interest for studying and working in S&T-related fields decrease if they have not developed affinity with S&T by the end of primary education (Turner & Ireson, 2010). The integration of S&T education into the primary school curriculum is thus widely supported (OECD, 2015).

In the last two decades, the Dutch government has therefore promoted the widespread implementation of science and technology (S&T) teaching in primary education. To this end, various national projects were funded, among most noticeably, Verbreding Techniek Basisonderwijs (VTB) (in English: Broadening Technology Education in Primary Education) and VTB-Pro (extended focus on teacher professionalization) by Platform Betá Techniek (in English: Platform Beta Technology). Schools were provided financial support to purchase science lesson examples for teaching S&T. Teachers were also offered (subsidized) teacher professionalization to acquire competency in teaching S&T. Although both projects generated significant interest among many schools, the project outcomes were limited (De Vries, Van Keulen, Peters, & Walma van der Molen, 2011). Evaluations indicated that relatively few primary teachers opted for the teacher professionalization (about 10 percent of all primary teachers in The Netherlands). In many Western countries, including The Netherlands, many primary teachers feel incapable of teaching S&T and thus often shy away from such teaching (Osborne & Dillon, 2008). In addition, the project outcomes indicated that few teachers appeared successful in involving other members of their school (including their school principals) with the school-wide implementation of S&T education. In turn, they received little organizational support

and guidance from their school leadership to reform their own practices, such as opportunities for teacher collaboration and classroom experimentation.

In response to the above project outcomes, in the year 2012, the Council for Primary Education and Platform Bèta Techniek established a special Exploratory Committee comprised of various Dutch education experts (Verkenningscommissie, 2013). The committee was charged with the task to recommend evidence-informed guidelines for an improved and sustained implementation of S&T teaching in Dutch primary education. International research proposed various guidelines to successfully introduce S&T education in primary schools (e.g., Lucas, Claxton, & Spencer, 2013; Osborne & Dillon, 2008; Pellegrino & Hilton, 2012). Based on this research, the Exploratory Committee proposed three main guidelines.

The first guideline concerned a broader meaning of 'S&T education' as a foundation for (pre-service) teacher professionalization, teaching materials, and the assessment of school development on this front. According to the Exploratory Committee, the concept of 'S&T' should be understood in broader terms than as a knowledge domain but rather as the process of inquiry, specifically, as 'a way of looking at the world' (Verkenningscommissie, 2013, p. 6). As such, 'S&T-minded' pupils are primarily characterized by their inquisitive attitudes, and by various higher-order thinking skills, such as creativity, critical thinking, and problem-solving. The committee stated that S&T education should not be understood as a separate school subject area, but rather as a transdisciplinary form of inquiry-focused teaching and learning linked to all school subject areas and lesson activities. This integral nature of S&T education would give pupils a better understanding of the interrelatedness of S&T with other school topics (such as literacy, geography, history and arts) and at the same time offer primary teachers increased opportunity to teach S&T. Such broader notions of primary S&T education are also proposed in the international literature (e.g., Claxton, 2007).

Secondly, based on the aforementioned project outcomes, the Exploratory Committee found that teachers play an essential role in the successful implementation of S&T education. Therefore, educational support should include teacher professionalization aimed at helping teachers acquire the knowledge, skills, and positive attitudes to teach S&T by inquiry and to alter their practices accordingly, such as by creatively 'infusing' inquiry teaching methodology into their usual teaching practices (Wilson, 2013). Moreover, the duration of such professionalization should be extended to offer teachers increased opportunity to apply what they have learned to their own classroom practice (Desimone, 2009).

Finally, the Exploratory Committee advised that primary schools themselves should learn to creative favorable conditions for teachers to implement and sustain S&T education in their classrooms. School principals with a strong inquiry-oriented vision about pupils' education, who prioritize school norms about classroom experimentation and teacher collaboration will likely be more successful with implementing S&T education than school principals who do not (Thurlings, Evers, & Vermeulen, 2015). The Exploratory Committee therefore called for school-wide approaches to the implementation of S&T education, in which complete school staffs, including the school principals, are collectively trained.

Goals and focus of the dissertation

The report offered by the Exploratory Committee prompted the start of the current doctoral research project in the year 2013 (funded by TechYourFuture, center of expertise in technology education). The present study mainly concerned the evaluation of a two-year school improvement program in which the complete school staffs of six Dutch primary schools were trained to integrate inquiry-focused (S&T) teaching into daily school practice. In the development of the program, the above-described main guidelines of the Exploratory Committee (Verkenningscommissie, 2013) were adopted. As such, the central goals of the program were explicitly linked to recent (inter)national policy on primary science and technology education. Descriptions of longitudinal and experimental intervention studies on the (school-wide) implementation of S&T education in primary schools are very relevant, but have so far been scarce. The present study aimed to help fill this void in the literature.

A main objective of the program was to positively develop the personal and professional attitudes of the participating teachers with concern to (teaching) S&T. This attitudinal approach adheres to the notions put forth by the Exploratory Committee (Verkenningscommissie, 2013) that teachers play an integral role in determining the successful adoption of S&T education. Based on attitude theory (Ajzen, 2001), it is assumed that teachers' attitudes towards (teaching) S&T determine the extent to which they teach S&T in their classroom (Osborne, Simon & Collins, 2003; Van Aalderen-Smeets, Walma van der Molen & Asma, 2012). A recent experimental study by our research group showed that in-service primary teachers' attitudes can be improved by six months of attitude-focused professional training (Van Aalderen-Smeets & Walma van der Molen, 2015). We thus placed this training at the basis of the intervention.

In addition, we developed a consecutive three-month training course that aimed to help participants familiarize with the broader concept of S&T teaching and learning, as formulated by the Exploratory Committee (Verkenningscommissie, 2013). It also aimed at helping participants acquire the didactic knowledge, skills, and positive attitudes to foster pupils' curiosity, inquiry ability beliefs, effort beliefs, achievement goal motivations, and higher-order thinking as an integral component of their daily teaching practices. The participating school principals did not receive specialized leadership training, but were required to participate in all course meetings and were encouraged to facilitate teachers' adoption of inquiry-based pedagogy during the program. In line with the recommendations of the Exploratory Committee, all teachers and the school principals of the participating schools collectively participated in both parts of the training course.

The effects of the professionalization treatment were subsequently examined (longitudinally) over a two-year time span at the level of the school principal, the teachers and the pupils in the middle and upper grades (grades 4, 5, and 6). These effects were compared to a control group of schools that did not receive the same treatment. A mixed-method approach to data collection was used to measure teachers' attitudes towards inquiry teaching, their inquiry teaching behavior, and school principals' leadership on teachers' adoption of inquiry-based pedagogy (e.g., Likert-type questionnaires and semi-structured interviews). At the pupil level, Likert-type questionnaires were used to examine changes in pupils' inquisitive, confident, and goal-driven thinking as a result of their teachers' changed inquiry teaching behavior. Taken together, these variables are proposed in the scientific literature as generally important outcome measures of S&T-focused school development.

However, preliminary research by the doctoral candidate revealed a scarcity of research on the nature and dimensions of pupils' curiosity within educational settings (see also Grossnickle, 2016). Within this research field, guidelines and instruments to encourage and assess pupils' curiosity in school barely exist and, when available, often poorly suit teachers' needs. This scarcity is somewhat surprising, as the encouragement of pupils' curiosity is widely promoted as one of the main objectives of S&T teaching in primary education, such as formulated by the Dutch Exploratory Committee. A few explorative studies in other countries suggest that pupils are offered litter opportunity to be inquisitive in primary school, including during S&T-focused lesson activities (Engel, 2015). However, this has never been properly investigated in The Netherlands.

As part of this dissertation, we thus also focused on the operationalization and measurement of pupils' curiosity in the context of primary S&T education. Based on attitude theory, we propose that pupils' attitudes towards (epistemic) curiosity precede their potential curiosity behavior in the classroom. This attitudinal perspective on pupils' curiosity and wonderment is closely related to the proposed interpretation of 'S&T-minded' pupils by the Exploratory Committee (Verkenningscommissie, 2013). In line with attitude theory (Ajzen, 2001), we advocate that a fruitful approach to stimulating pupils' epistemic wonderment, questions, and ideas in the classroom is to foster their epistemic images of curiosity and their positive beliefs and feelings about the value, social appreciation, and opportunity of being inquisitive learners. Based on this approach, we developed and validated a questionnaire that can be used to reliably and efficiently measure pupils' images of and attitudes towards curiosity among middle and upper-level pupils, which we call the Children's Images of and Attitudes towards Curiosity (CIAC) questionnaire. In addition, we developed didactic guidelines for primary teachers to positively develop these images and attitudes among their pupils. All participating teachers were trained during the professional training course to adopt these guidelines as explicit lesson objectives of their daily classroom practices. We also used the CIAC questionnaire to measure (changes in) pupils' attitudes towards curiosity as part of the aforementioned longitudinal effect study.

Overview of the studies of the dissertation

The second half of this dissertation mainly describes the evaluation of the above school development program. As previously described, however, preliminary research indicated a scarcity of research on the nature and dimensions of pupils' curiosity in the school context. Because pupils' curiosity was deemed a main objective of primary S&T education, and should therefore be included as an explicit measure for assessing inquiry-focused school improvement in this context, separate theoretical and empirical research had to be conducted first. Therefore, the dissertation comprises two consecutive parts. The first two studies of the dissertation focus on the operationalization and measurement of pupils' curiosity. The last two studies focus on the evaluation of the school development program. Below, we summarize the main findings of the four studies.

Study 1: Do children express curiosity at school? Exploring children's experiences of curiosity inside and outside the school context

Despite the global emphasis on the stimulation of pupils' curiosity in primary education, little is known about pupils' *own* concepts about, feelings towards, and experiences with 'being curious' in school. Such insight, however, serves to determine what aspects of pupils' curiosity require teachers' attention and, as such, serve as an important stepping stone for the development of curiosity-focused pedagogy, teacher professionalization, and measurement instruments.

In the first study of this dissertation (chapter 2), we thus set out to explore primary pupils' pre-existing concepts of, feelings towards and experiences with various types of 'curiosity' from all grade levels of two Dutch primary schools. The results of the study indicated that most pupils barely understood curiosity in epistemic terms, but that they associated the term predominantly with something that belongs to the social domain (e.g., eavesdropping, prying, gossiping). Partly because of this narrow conception, the pupils did not attribute much learning-value to being curious in school. In addition, they felt generally discouraged by their teachers to express their epistemic questions and ideas in class. However, many pupils reported to be actively curious about a diverse range of complex science topics outside of the school context, such as at home or during holiday trips. This result is in line with studies in the curiosity literature that pupils are naturally curious.

Most of pupils' teachers in the study indicated to feel concerned about pupils' limited curiosity experiences in school, as the teachers all perceived the clear educational value of pupils' curiosity. Many teachers indicated to perceive pupils' curiosity as a natural "by-product" of their inquiry-based lesson activities and that its stimulation did thus not require their explicit attention. The results of this study suggest that this perception is likely false and that, therefore, teachers *should* explicitly foster pupils' epistemic curiosity as part of their regular lesson objectives, including inquiry-oriented lesson activities.

This study was published as: Post, T., & Walma van der Molen, J. H. (2018). Do children express curiosity at school? Exploring children's experiences of curiosity inside and outside the school context. *Learning, Culture and Social Interaction, 18*, 60–71.

Study 2: Development and validation of a questionnaire to measure primary school children's images of and attitudes towards curiosity (the CIAC questionnaire)

The findings of the first study led us to consider the aspects of pupils' curiosity that would require particular attention by primary teachers. These considerations are described as part of the second study of this dissertation (chapter 3).

Rather than focusing on pupils' curiosity behavior, skills, traits, or states, we propose to shift focus towards the positive development and assessment of pupils' *perceptions* of curiosity. Such perceptions can be understood in terms of pupils' images of and attitudes towards curiosity. A large body of research shows that attitudinal and motivational beliefs affect behavior, such as the types of emotions pupils experience during learning and the types of learning strategies they employ. Based on attitude and curiosity theory, we thus advocate that a fruitful approach to stimulating pupils' epistemic wonderment, questions, and ideas in the classroom is to foster their epistemic images of curiosity and their positive beliefs and feelings about being inquisitive learners. This attitudinal perspective on the stimulation of pupils' curiosity closely matches the proposed definition of 'S&T-minded' pupils by the Dutch Exploratory Committee (see chapter 1).

However, because this perspective was new, no corresponding operationalizations and measurement instrument yet existed. Moreover, there are hardly any clear definitions nor validated measures of pupils' curiosity within educational settings. As part of the second study (chapter 3), we thus made use of attitude, motivation, and curiosity theory to propose relevant components of pupils' images of and attitudes towards curiosity. We translated these theoretical image and attitude components into corresponding measurement scales and validated these scales by means of qualitative and extensive quantitative methods. We call the instrument the Children's Images of and Attitudes towards Curiosity (CIAC) questionnaire and is suited for pupils in the 4th, 5th, and 6th grades. In line with attitude theory, we also provide evidence for the predictive power of pupils' images and attitudes on their achievement-related motivations.

In sum, the second study of this dissertation thus contributes to curiosity research by unpacking the overarching concept of pupils' curiosity in its underlying image and attitude components. It thereby, also, offers a novel and relevant perspective on the stimulation and assessment of pupils' curiosity in primary school, namely by positively developing pupils' images of and attitudes towards epistemic curiosity. To that end, primary teachers should foster positive classroom cultures in

which pupils (learn to) perceive the educational value of being inquisitive learners, derive a sense of pleasure and pride from asking inquisitive questions or wondering about alternative answers to questions during lesson activities, and feel that their teachers and peers encourage such behavior.

This study was published as: Post, T., & Walma van der Molen, J. H. (2018). Development and validation of a questionnaire to measure primary school children's images of and attitudes towards curiosity (the CIAC questionnaire). *Motivation and Emotion, 43,* 159–178.

This paper received the PhD Best Paper Award in the year 2017 by the Motivation and Emotion special interest group of the European Association for Research on Learning and Instruction (EARLI) conference in Finland.

Study 3: Effects of a longitudinal school development program on primary teachers' attitudes towards inquiry teaching and their inquiry teaching practices

In the third study of this dissertation, the complete staffs of six Dutch primary schools were trained to integrate S&T education into daily school practice. A quasi-experimental research design was used with pretest and posttest measurements to investigate the effects of professionalization treatment on teachers' attitudes towards inquiry teaching and their teaching behavior in comparison to a control group of schools that did not receive this type of training. This control group later received the same professional training (i.e., delayed intervention).

Differences in the effects of the professional training on teachers' attitudes and behavior were examined based on the possible differences in school leadership for S&T education between the participating schools. Therefore, the program effects were examined on the treatment level (i.e., across the treatment schools compared to the non-treatment schools) and at the individual school level. Quantitative data were collected by means of questionnaires collected from school principals and teachers. In addition, the school principals, teachers and pupils were also interviewed.

Results indicate that the program significantly improved teachers' attitudes and inquiry teaching behavior of the first treatment group and that teachers' improved practices persisted to one year after training. Many of the participating teachers and school principals stated at the end of the research project that the program met their expectations and professionalization needs. However, our findings highlight the challenge that primary teachers may face when aiming to stimulate pupils' creative,

higher-order thinking. In addition, school leadership remained largely moderate throughout the program and appeared to explain little variation in teachers' inquiry teaching development among the participating schools. Also, the same training showed a limited effect in the case of the delayed intervention group, but this limitation seemed to be largely caused by incidental and school-specific factors that hindered school improvement (e.g., mismatch of training needs, being discontent about the personal style of one of the trainers, the introduction of new lesson methods).

The findings of the study thus suggest that a combination of extensive attitude-focused and didactical training may provide primary school teams the (minimal) preparatory training to integrate S&T education in daily school practice. We note that primary school teams, even those who are highly committed to adopt such practice, do not necessarily possess positive attitudes towards inquiry teaching to begin with. Attitude training should thus be regarded a vital and explicit starting point of inquiry-focused teacher professionalization. Attitude change and school practice reform do not happen overnight and, as such, our findings suggest that inquiry-focused school development requires substantial teacher professional development (i.e., taking at least one school year of training). Although schools likely prefer to participate in few, short, and hands-on workshops, our findings suggest that such workshops will likely affect little significant change in what teachers think and do. Lastly, we show the importance of a 'school-wide' approach to the implementation of inquiry-based pedagogy in primary schools, because such an approach stimulates teacher collaboration and the cultivation of shared positive norms on practice reform in schools.

This study is accepted for publication as: Post, T., & Walma van der Molen, J. H. (in press). Effects of a longitudinal school development program on primary teachers' attitudes towards inquiry teaching and their inquiry teaching practices. *Teacher College Records.*

Study 4: Effects of an inquiry-focused school improvement program on the development of pupils' attitudes towards curiosity, their implicit ability and effort beliefs, and goal orientations

Based on the findings of the third study, we subsequently examined the effects of the improved teaching behavior of the middle and upper school teachers on the subsequent improvements in their pupils' attitudes to curiosity, their implicit views

on the malleability of their inquiry ability, their implicit opinions about the influence of their efforts on their learning performance, and their achievement-related motivations for inquiry. These variables are proposed in the literature as relevant outcome measures of pupil development in the field of S&T in primary education.

Therefore, quantitative data were collected from all 4th, 5th, and 6th grade pupils before, immediately after, and one school year after their teachers' professionalization treatment. The effects of the treatment on the development of the pupils were investigated on the basis of a longitudinal, experimental pre-test post-test study design with a control group. A Structural Equation Modeling approach was first used to examine the relationships among pupils' attitudes, beliefs, and motivations. In line with attitude and motivation theory, the results of the study show that pupils' attitudes towards curiosity and their implicit ability beliefs predicted their efforts beliefs and goal orientation motivations. In addition, the intervention affected small but significant improvements in pupils' attitude, belief and motivation scores during the two years of the program.

The findings of study 4 thus provide additional evidence to support the effectiveness of the professionalization treatment, namely by showing that teachers' improved inquiry teaching behavior affected subsequent improvements in pupils' inquiry-related attitudes, beliefs, and motivations. Furthermore, with this study, we show that pupils' attitudes towards curiosity and their implicit ability beliefs likely *affect* their efforts and motivations to be inquiry-minded learners. Thereby, we provide novel insight into the psychological mechanism that may underlie pupils' potential inquiry behavior in the classroom. The stimulation of positive attitudes towards epistemic curiosity and implicit ability beliefs may thus be a valuable supplement to typical forms of inquiry-based pedagogy.

If we want to be successful in helping pupils mature into inquisitive, creative, and confident thinkers, such as promoted by twenty-first century (S&T) education standards, we advocate that teachers should thus make pupils aware of *why* 'being inquiry-minded' in school is fruitful in the first place. This means that, in the midst of literacy, history, mathematics and science lessons, teachers should explicitly cultivate pupils' positive perceptions about the educational value of asking epistemic questions, a sense of pleasure and pride from seeking creative solutions to problems, and the confidence that their repeated attempts to face and tackle such problems boosts their potential inquiry ability. As shown in the present study, teachers can make pupils aware of their (implicit) beliefs about epistemic curiosity and inquiry ability and,

subsequently, challenge them to adopt more positive beliefs about these matters by reflection and discussion.

Our findings may thereby inform the further development of school development projects aimed at stimulating pupils' inquiry in primary education, namely by fostering pupils' attitudes towards curiosity and their implicit ability beliefs.

This study is submitted for publication as: Post, T., & Walma van der Molen, J. H. (2019). Effects of an inquiry-focused school improvement program on the development of pupils' attitudes towards curiosity, their implicit ability and effort beliefs, and goal orientations. Manuscript submitted for publication.

Conclusions

The main findings of the present study indicate that a school-wide approach to the implementation of S&T education, coupled with extensive attitude-focused and didactical professional training that adheres to a broader focus on S&T education, can help primary schools integrate inquiry-focused (S&T) teaching practice in their school organization. This dissertation thereby provides empirical evidence to support the efficacy of the main guidelines proposed by the Exploratory Committee (Verkenningscommissie, 2013) for the successful implementation of S&T teaching practice in primary education.

At the same time, our findings highlight that it is not easy for primary teachers to implement such teaching practice on their own. Changing habits is difficult, especially   when change requires the mastery of new complex teaching skills. In our view, scientific progress in this field can only be achieved when researchers are aware of the complexity of S&T-focused interventions. They should make explicit and substantiated decisions about what factors to foster and measure, at multiple school levels, and use evidence-based interventions and methodologically sound instruments wherever possible.

However, other important challenges remain. For inquiry-focused interventions to move beyond isolated and promising examples toward more widespread education change, larger systemic issues and policies will need to be considered. For example, when the positive development of pupils' images of and attitudes towards curiosity would be incorporated into national primary education standards (as advocated in this dissertation), this would provide a strong incentive for

school districts and teachers to emphasize its development and assessment as part of school practice. Foundations and federal officials should therefore promote research aimed at more clearly defining and developing assessments of such pupil qualities, initially for research purposes, and later for summative and formative assessment (see also Geisinger, 2016; Inspectie van het Onderwijs, 2019).

Only when pupils' inquiry is perceived by primary teachers as integral to school learning, no less so than solving a mathematical problem or writing an essay, will we likely see significant changes happen in classroom practice. Teachers need to become convinced about the use, pleasure, and opportunity of learning by inquiry themselves, which will empower them to act as genuine inquiry-driven role-models to their pupils. In turn, they should cultivate in pupils a sense that school is not only about memorizing historical facts, understanding how to gather information from the internet, or applying arithmetic to calculate surface areas. Rather, that school is ultimately about becoming an inquisitive explorer, an original thinker, and a resilient competitor in whatever future field they may find themselves. Teachers have to explain to pupils why their teachers will keep challenging them with novelty and complexity during their academic careers, what they have to put in to excel in life, and what they can expect to gain if they truly commit themselves.

This way, school can become a place where pupils' inquisitive, creative and bold thinking not only survives, but flourishes.

Samenvatting

Nederlandse samenvatting (Dutch summary)

Introductie

Dit proefschrift beschrijft de evaluatie van een tweejarig schoolontwikkelingsproject in het basisonderwijs, waarin zes voltallige schoolteams werden geprofessionaliseerd voor het integreren van W&T-onderwijs (een verzamelterm voor onderwijs op het gebied van wetenschap en technologie) in hun dagelijkse lespraktijk. Het belang van W&T-onderwijs neemt wereldwijd toe, omdat wetenschappelijke en technologische innovaties onze samenleving op alle niveaus sneller en ingrijpender veranderen (World Economic Forum, 2018). Alle kinderen zullen dus op school vertrouwd moeten raken met deze innovaties om op te kunnen groeien tot volwaardige deelnemers van de maatschappij (National Research Council, 2012). Onderzoek suggereert dat de natuurlijke interesses van kinderen voor W&T-gerelateerde studies en beroepen bovendien afnemen als zij daar niet al in het basisonderwijs positief mee in aanraking komen (Turner & Ireson, 2010). De structurele verankering van W&T in het basisonderwijs wordt daarom breed gedragen (OECD, 2015).

In de laatste twee decennia werden daartoe in Nederland verschillende landelijke projecten met overheidsmiddelen opgezet, zoals in het bijzonder de programma's Verbreding Techniek Basisonderwijs (VTB) en VTB-Pro (aangevuld met leerkrachtprofessionalisering). Daarbij ontvingen basisscholen financiële middelen om uitgewerkte lesvoorbeelden op het gebied van W&T aan te schaffen. Tevens werd (gesubsidieerde) leerkrachtprofessionalisering aangeboden om leraren bekend te maken met het onderwijzen van W&T. Hoewel deze inspanningen zorgden voor een opleving van W&T in het scholenveld, bleef het gewenste eindresultaat beperkt (De Vries, Van Keulen, Peters, & Walma van der Molen, 2011). Resultaten toonden dat relatief weinig leraren kozen om deel te nemen aan de leerkrachtprofessionalisering (ongeveer 10 procent van alle basisschoolleraren in Nederland). In veel Westerse landen, waaronder Nederland, voelen veel basisschoolleraren zich onzeker en oncomfortabel bij het onderwijzen van W&T en vermijden zij zulk onderwijs daarom het liefst (Osborne & Dillon, 2008). Resultaten lieten ook zien dat het deel geprofessionaliseerde leraren door VTB-Pro er nauwelijks in slaagde collega's (inclusief hun schoolleiders) te betrekken in de school-brede implementatie van W&T. Leraren ontvingen weinig organisatorische ondersteuning en begeleiding van

hun schoolleiding bij hun onderwijsvernieuwing in de klas, zoals mogelijkheden tot samenwerken en experimenteren met collega's.

De bovenstaande projectuitkomsten waren in het jaar 2012 aanleiding voor de PO-raad en Platform Bèta Techniek een Verkenningscommissie 'Wetenschap en Technologie Primair Onderwijs' in te stellen met de opdracht advies te geven voor een duurzame implementatie van W&T in het Nederlandse basisonderwijs (Verkenningscommissie, 2013). Internationaal wetenschappelijk onderzoek bood inmiddels verschillende richtlijnen voor de succesvolle implementatie van W&T in het basisonderwijs (e.g., Lucas, Claxton, & Spencer, 2013; Osborne & Dillon, 2008; Pellegrino & Hilton, 2012). Op basis van dit (inter-)nationale onderzoek formuleerde de Verkenningscommissie drie hoofdadviezen.

Het eerste advies betrof een betere begripsbepaling van W&T-onderwijs als uitgangspunt voor (na-)scholing voor basisschoolleraren, lesmaterialen en de beoordeling van schoolprestaties op dit vlak. Het begrip 'wetenschap en technologie' zou volgens de commissie breder moeten worden opgevat dan een kennisdomein, maar als 'een manier van kijken naar de wereld' (Verkenningscommissie, 2013, p. 6). Deze kijkwijze wordt in de basis gekenmerkt door een nieuwsgierige, onderzoekende houding, waarbij verschillende hogere-orde denkvaardigheden van belang zijn, zoals creatief, kritisch en probleemoplossend denken. De Verkenningscommissie stelde dat W&T-onderwijs als zodanig geen apart schoolvak is, maar een vakoverstijgende vorm van onderzoekend en ontwerpend leren, gekoppeld aan alle schoolvakken en lesactiviteiten. Dit integrale karakter van W&T-onderwijs geeft kinderen meer inzicht in de verwevenheid van W&T met reguliere vakgebieden (zoals met taal, aardrijkskunde, geschiedenis en kunst) en biedt leraren tegelijkertijd meer gelegenheid voor het onderwijzen van W&T. Soortgelijke opvattingen van W&T-onderwijs worden ook internationaal voorgesteld (zie bijvoorbeeld Claxton, 2007).

Ten tweede constateerde de Verkenningscommissie dat de rol van de leraar bij de implementatie van W&T-onderwijs doorslaggevend bleek in de eerdergenoemde projecten. Het aanbod van (na-)scholing op het gebied van W&T-onderwijs zou daarom nadrukkelijker aandacht moeten schenken aan de kennis, vaardigheden en attitudes die leraren nodig hebben voor het implementeren van W&T-onderwijs, zoals voor vakoverstijgend, onderzoekend en ontwerpend leren en voor het flexibel omgaan met lesmethodes (Wilson, 2013). De opleidingsduur zou bovendien moeten worden verlengd, zodat leraren meer gelegenheid krijgen om het geleerde te bestendigen in de eigen klaspraktijk (Desimone, 2009).

Tot slot, adviseerde de Verkenningscommissie dat basisscholen daarbij zélf gunstige voorwaarden moeten leren scheppen voor leraren om met W&T in de eigen klas aan de slag te gaan. Scholen met een expliciete visie op W&T in hun onderwijsbeleid, die leraren ruimte geven voor onderlinge samenwerking, en een positieve schoolcultuur stimuleren ten aanzien van het (leren) onderwijzen van W&T, bleken over het algemeen beter in staat om W&T-onderwijs te implementeren (Thurlings, Evers, & Vermeulen, 2015). De Verkenningscommissie pleitte daarom voor een 'schoolbrede' invoering van W&T-onderwijs, waarbij voltallige schoolteams, inclusief hun schoolleiders, gezamenlijk nageschoold worden.

Doelen en focus van het promotieonderzoek

Bovenstaande ontwikkelingen waren in het jaar 2013 aanleiding voor de start van een promotieonderzoek (gefinancierd door TechYourFuture, expertisecentrum voor onderwijs bèta en technologie). Het onderzoek betrof hoofdzakelijk de evaluatie van een tweejarig schoolontwikkelingsproject waarin zes voltallige schoolteams werden nageschoold voor het verankeren van W&T-onderwijs in de eigen schoolorganisatie. Bij de ontwikkeling van het programma zijn de bovenstaande hoofdadviezen van de Verkenningscommissie (2013) opgenomen. Als zodanig relateerden de doelen van het programma expliciet aan (inter-)nationaal onderwijsbeleid op het gebied van W&T-onderwijs. Beschrijvingen van longitudinale en experimentele interventiestudies naar de (schoolbrede) inbedding van W&T-onderwijs in basisscholen zijn momenteel zeer relevant, maar nog altijd schaars. Het huidige promotieonderzoek vult hiermee een leemte in de internationale wetenschappelijke literatuur.

Een centrale doelstelling van het schoolontwikkelingsproject is geweest om de persoonlijke en professionele attitudes van de deelnemende leraren ten aanzien van het (onderwijzen van) W&T positief te ontwikkelen. Deze aanpak sluit nauw aan bij de constatering van de Verkenningscommissie (2013) dat de rol van leraren doorslaggevend is bij de implementatie van W&T-onderwijs. Op basis van attitudetheorie (Ajzen, 2001) wordt verondersteld dat deze opvattingen van leraren bepalend zijn voor de mate waarin zij W&T onderwijzen in de klas (Osborne, Simon & Collins, 2003; Van Aalderen-Smeets, Walma van der Molen, & Asma, 2012). In een recente experimentele studie door onze onderzoeksgroep bleek een zelfontwikkelde, attitude-gerichte nascholing van zes maanden effectief in het verbeteren van de attitudes van de deelnemende leraren en bij hun implementatie van W&T-onderwijs in de klas (Van Aalderen-Smeets & Walma van der Molen, 2015). Deze attitude-

gerichte nascholing vormde daarom de basis van het huidige schoolontwikkelingsproject.

Daarnaast is een aanvullende nascholing (van drie maanden) ontwikkeld die zich nadrukkelijker richtte op de bredere opvatting van W&T-educatie, zoals geformuleerd door de Verkenningscommissie (2013). Deze aanvulling beoogde de didactische kennis en vaardigheden van de deelnemende leraren te verbeteren voor het koppelen van onderzoekend en ontwerpend leren aan bestaande lesinhouden en het creatiever omgaan met lesmethodes. De deelnemende schoolleiders ontvingen geen gespecialiseerde leiderschapstraining, maar moesten wel alle cursusbijeenkomsten bijwonen en de onderwijsvernieuwing van de leerkrachten faciliteren tijdens het programma. In lijn met de adviezen van de Verkenningscommissie, namen álle leraren en de schoolleiders van de deelnemende scholen gezamenlijk deel aan beide delen van het nascholingsprogramma.

De effecten van de nascholing werden vervolgens gedurende twee jaar (longitudinaal) onderzocht op het niveau van de schoolleider, de leraren en de leerlingen in de midden- en bovenbouwgroepen (groepen 6, 7 en 8). Deze effecten werden vergeleken met een controlegroep van scholen waar niet aan dit soort schoolontwikkeling werd gedaan. Daarvoor is gebruikt gemaakt van verschillende kwantitatieve en kwalitatieve onderzoeksmethoden. Bij de schoolleiders is hiervoor gekeken naar mogelijke veranderingen in hun schoolbeleid, het faciliteren van onderlinge samenwerking tussen leraren, en het stimuleren van een positieve schoolcultuur ten aanzien van het (leren) onderwijzen van W&T op de eigen school. Bij de leraren is gekeken naar veranderingen in hun professionele attitudes ten aanzien van het onderwijzen van W&T, veranderingen in hun attitudes ten aanzien van eigen lesontwerp en de (voortvloeiende) veranderingen in hun dagelijkse W&T-onderwijs. Tot slot is bij de leerlingen gekeken naar veranderingen in hun nieuwsgierige, zelfverzekerde en prestatiegerichte denken over tijd, als resultaat van het verbeterde lesgedrag van hun leraren. Samengenomen, worden deze variabelen in de wetenschappelijke literatuur vaak voorgedragen als belangrijke doelmaten van W&T-gerichte schoolontwikkeling.

Uit vooronderzoek door de promovendus bleek echter dat tot nu toe nauwelijks duidelijkheid bestaat in de literatuur over de aard en de dimensies van de 'nieuwsgierigheid' van basisschoolleerlingen als motor voor onderzoekend en ontwerpend leren. Dit is opmerkelijk, omdat onderwijsbeleid en onderzoek nieuwsgierigheid bij kinderen als de basis zien voor W&T en het onderzoekende en ontwerpende leren dat daarbij hoort, zoals ook de Verkenningscommissie (2013).

Binnen dit onderzoeksveld wordt echter nog weinig gewerkt met gestandaardiseerde instrumenten en daar waar deze zijn ontwikkeld, sluiten deze tot nu toe slecht aan op de behoeften. Enkele exploratieve studies in het buitenland suggereren dat het nieuwsgierige, onderzoekende denken van kinderen bovendien niet veel voorkomt in de basisschool, zelfs niet tijdens lesactiviteiten gericht op W&T (zie Engel, 2015). Dit is in Nederland echter nog nooit goed onderzocht.

In dit promotieonderzoek is daarom ook aandacht geschonken aan de operationalisering en meting van de nieuwsgierigheid van basisschoolleerlingen in het kader van W&T-onderwijs. Op basis van attitudetheorie stellen we voor dat de *attitudes* van kinderen ten aanzien van (epistemische) nieuwsgierigheid voorafgaan aan hun potentiële nieuwsgierige, onderzoekende gedrag in de klas. Dit perspectief op 'nieuwsgierigheid' sluit nauw aan bij de begripsbepaling van W&T zoals voorgesteld door de Verkenningscommissie (2013), die nieuwsgierigheid ook definieerde in termen van een houding. In lijn met attitudetheorie (Ajzen, 2001), veronderstellen we dat het nieuwsgierige onderzoekende denken van kinderen in de klas pas kan ontstaan (en groeien) wanneer kinderen zélf de waarde inzien van zulk denken voor hun eigen leerprestaties op school, de waarde inzien van nieuwsgierige onderzoekende mensen voor kennisontwikkeling en innovatie in algemene zin, sociale waardering ervaren voor hun nieuwsgierige ideeën en vragen in de klas, en zich bekwaam voelen in nieuwsgierig, onderzoekend denken. We ontwikkelden en valideerden daarom een vragenlijst waarmee de bovenstaande attitude-componenten onder midden- en bovenbouwleerlingen gemeten kunnen worden (genaamd de Children's Images of and Attitudes towards Curiosity (CIAC) vragenlijst). Daarbij hebben we didactische richtlijnen voor leraren opgesteld voor het positief ontwikkelen van deze attitude-componenten onder leerlingen. Alle deelnemende leraren werden tijdens het nascholingsprogramma getraind om deze richtlijnen als expliciete lesdoelen te integreren in hun dagelijkse lespraktijk. De CIAC-vragenlijst is ingezet om (veranderingen in) de attitudes van basisschoolleerlingen ten aanzien van nieuwsgierigheid te meten als onderdeel van de eerdergenoemde longitudinale effectstudie.

Overzicht van de deelstudies van het proefschrift

Het tweede deel van het proefschrift beschrijft de evaluatie van het bovenstaande schoolontwikkelingsprogramma. Zoals reeds beschreven, bleek echter uit vooronderzoek dat er in de literatuur nog nauwelijks duidelijkheid bestaat over het

nieuwsgierige, onderzoekende denken van kinderen op de basisschool. Omdat zulk denken bij kinderen centraal staat in de doelstellingen van W&T-onderwijs, en dus als een expliciete beoordelingsmaat zou moeten worden meegenomen voor het inschatten van schoolprestaties op dit vlak, moest hier eerst afzonderlijk theoretisch en empirisch onderzoek naar worden verricht. Het proefschrift bestaat daarom uit twee opvolgende delen. De eerste twee studies van het promotieonderzoek richten zich op de verkenning, operationalisering en meting van nieuwsgierigheid bij kinderen. De laatste twee studies richten zich op de evaluatie van het schoolontwikkelingsprogramma. Hieronder volgt een beknopte samenvatting van de vier deelstudies.

Studie 1: De percepties van basisschoolleerlingen over hun eigen ervaringen met nieuwsgierig, onderzoekend denken binnen en buiten school

Ondanks de wereldwijde aandacht voor het stimuleren van het nieuwsgierige, onderzoekende denken van leerlingen in het basisonderwijs, is er nog weinig bekend over de eigen percepties en ervaringen van leerlingen met epistemische nieuwsgierigheid op de basisschool. Zulk inzicht helpt echter bij het bepalen van welke aspecten van de nieuwsgierigheid van leerlingen aandacht behoeven van leraren en dient als zodanig als een belangrijk aanknopingspunt voor de ontwikkeling van passende pedagogiek, leerkrachtprofessionalisering en meetinstrumenten.

In de eerste studie van dit proefschrift (hoofdstuk 2), interviewden we daarvoor eerst een honderdtal basisschoolleerlingen van alle groepen van de basisschool over hun percepties van, gevoelens bij en ervaringen met verschillende typen nieuwsgierigheid binnen en buiten de schoolcontext. Uitkomsten van de studie suggereren dat kinderen 'nieuwsgierigheid' voornamelijk associëren met sociaal gedrag (bijv. afluisteren en roddelen). Deels door deze smalle opvatting, lijken kinderen weinig belang te zien bij het ontwikkelen van hun nieuwsgierigheid op school en voelen zij zich door hun leraren veelal ontmoedigd om hun nieuwsgierige ideeën en vragen over de lesstof in de klas te delen. De kinderen bleken echter wel nieuwsgierig naar een scala aan ingewikkelde, wetenschappelijke onderwerpen buiten school, zoals thuis en tijdens vakantie. Deze bevinding is in lijn met onderzoek naar nieuwsgierigheid dat uitwijst dat kinderen van nature nieuwsgierig zijn.

De meeste leraren van de leerlingen gaven aan bezorgd te zijn over de beperkte ervaring van hun leerlingen met nieuwsgierig-zijn op school, omdat zij allen de educatieve waarde van de nieuwsgierigheid van kinderen zagen. Veel van de leraren

gaven aan de nieuwsgierigheid van de leerlingen te zien als een natuurlijk "bijproduct" van hun W&T-lessen en dat de stimulering ervan dus niet om hun expliciete aandacht zou vragen. De resultaten van deze studie suggereren dat deze perceptie waarschijnlijk niet klopt en dat leraren de nieuwsgierigheid van leerlingen juist wél expliciet moeten bevorderen als onderdeel van hun (W&T-gerichte) lesdoelen.

Studie 1 is gepubliceerd als: Post, T., & Walma van der Molen, J. H. (2018). Do children express curiosity at school? Exploring children's experiences of curiosity inside and outside the school context. *Learning, Culture and Social Interaction, 18*, 60–71.

Studie 2: De ontwikkeling en validatie van een meetinstrument voor de beelden en attitudes van basisschoolleerlingen ten aanzien van epistemische nieuwsgierigheid

De resultaten van de eerste studie waren aanleiding voor het bepalen van de aspecten van de nieuwsgierigheid van basisschoolleerlingen die met name aandacht van leraren behoeven. Deze bepaling wordt beschreven als onderdeel van de tweede studie van dit proefschrift (hoofdstuk 3). In plaats van een focus op het gedrag, de vaardigheid, de persoonskenmerken of de toestanden van de nieuwsgierigheid van leerlingen, stellen we voor de aandacht te verleggen naar de verbetering en beoordeling van de *percepties* van leerlingen ten aanzien van nieuwsgierigheid. Zulke percepties kunnen worden begrepen in termen van de *beelden* en *attitudes* van leerlingen ten aanzien van nieuwsgierigheid. Internationaal onderzoek wijst uit dat attitude gedrag beïnvloedt, zoals de typen emoties die leerlingen ervaren tijdens het leren en de soorten leerstrategieën die zij inzetten. Gebaseerd op attitude- en nieuwsgierigheidstheorie, stellen we daarom dat interventies gericht op het stimuleren van de nieuwsgierigheid van leerlingen expliciet aandacht moeten schenken aan het bijbrengen van 'epistemische' beelden van nieuwsgierigheid bij leerlingen en het verbeteren van hun overtuigingen van en gevoelens bij 'nieuwsgierig-zijn' in de klas. Deze attitude-gerichte aanpak sluit nauw aan bij het advies van de Verkenningscommissie, die 'W&T-georiënteerde' leerlingen ook hoofdzakelijk duidt in termen van hun attitudes ten aanzien van nieuwsgierigheid.

Omdat ons bovenstaande perspectief op de nieuwsgierigheid van leerlingen nog niet eerder bestond, bestonden er ook nog geen bijbehorende operationaliseringen en meetinstrumenten in de literatuur. Bovendien bestaan er nog nauwelijks duidelijke definities en gevalideerde meetschalen van de soort van nieuwsgierigheid van

kinderen die relevant is in de schoolcontext. Als onderdeel van de tweede studie, maakten we daarom gebruik van theorie over attitude, motivatie en nieuwsgierigheid om de belangrijkste componenten van de beelden en attitudes van basisschoolleerlingen ten aanzien van nieuwsgierigheid te definiëren. We vertaalden deze componenten vervolgens in bijbehorende meetschalen waarmee deze componenten onder midden- en bovenbouwleerlingen van de basisschool kunnen worden gemeten. Ook valideerden we deze schalen met behulp van kwalitatieve en uitgebreide kwantitatieve methoden. Deze schalen vormen een nieuw meetinstrument dat we de 'Children's Images of and Attitudes towards Curiosity' (CIAC) vragenlijst noemen. In lijn met attitudetheorie, tonen we in deze studie ook aan dat de beelden en attitudes van leerlingen voorspellend zijn voor hun prestatiemotivaties.

Samenvattend, draagt de tweede studie van dit proefschrift bij aan nieuwsgierigheidsonderzoek door het overkoepelende concept van de nieuwsgierigheid van leerlingen te ontrafelen in belangrijke, onderliggende beeld- en attitude-componenten. Deze studie biedt daarmee een nieuw en relevant perspectief op het stimuleren en meten van de nieuwsgierigheid van leerlingen op de basisschool, namelijk door het positief ontwikkelen van hun beelden en attitudes ten aanzien van nieuwsgierigheid. Daartoe moeten leraren een positief klasklimaat cultiveren, waarin leerlingen de waarde (leren) inzien van hun nieuwsgierige gedrag, gevoelens van plezier en trots ontlenen aan het stellen van nieuwsgierige vragen of het verwonderen over alternatieve antwoorden op vragen tijdens lesactiviteiten, en voelen dat hun leraren en klasgenoten zulk gedrag aanmoedigen.

Studie 2 is gepubliceerd als: Post, T., & Walma van der Molen, J. H. (2018). Development and validation of a questionnaire to measure primary school children's images of and attitudes towards curiosity (the CIAC questionnaire). *Motivation and Emotion, 43,* 159–178. Dit onderzoeksartikel ontving in het jaar 2017 de 'PhD Best Paper Award' van de internationale onderzoeksgroep Motivation & Emotion van de European Association for Research on Learning and Instruction (EARLI) conferentie in Finland.

Studie 3: De effecten van een langdurig schoolontwikkelingsproject op de attitudes van basisschoolleraren ten aanzien van het onderwijzen van onderzoekend leren in hun dagelijkse lespraktijk

In de derde studie van dit proefschrift werden alle leraren van de deelnemende basisscholen nageschoold om W&T-onderwijs te integreren in hun bestaande lespraktijk. Voor de evaluatie van deze nascholing werd een experimentele onderzoeksopzet gebruikt met voor- en nametingen om de effecten van de nascholing op de attitudes van leraren ten aanzien van het onderwijzen van W&T en hun lesgedrag te onderzoeken in vergelijking met een controlegroep van scholen waar niet aan dit soort schoolontwikkeling werd gedaan. Deze controlegroep ontving later dezelfde nascholing (uitgestelde interventie).

Verschillen in de effecten van de nascholing op de leraren werden verkend in relatie tot de mogelijke verschillen in schoolleiderschap voor W&T onder de deelnemende scholen. Daarvoor werden effecten dus op zowel het conditieniveau als op het niveau van de individuele scholen bekeken. Onderzoeksgegevens werden verzameld met behulp van vragenlijsten die onder de schoolleiders en de leraren werden afgenomen. Ook werden de schoolleiders, leraren en leerlingen geïnterviewd.

De uitkomsten van de studie tonen dat de nascholing significante verbeteringen teweegbracht in de attitudes en het lesgedrag van de leraren in de eerste interventiegroep en dat dit verbeterde lesgedrag voortduurde tot het einde van het opvolgende schooljaar. Echter bleef de mate van schoolleiderschap voor W&T op de scholen gematigd positief en verklaarde dit leiderschap weinig verschil in de ontwikkeling van de leraren gedurende het programma. Dezelfde nascholing had bij de uitgestelde interventiegroep maar beperkt effect, maar deze beperking leek grotendeels te zijn veroorzaakt door incidentele en school-specifieke factoren die schoolverbetering belemmerden (bijv. andere trainingsbehoeften, ontevredenheid over de persoonlijke stijl van een van de trainers, de introductie van nieuwe lesmethoden).

De resultaten van de studie suggereren dat een combinatie van uitgebreide attitude-gerichte en didactische training basisschoolleraren de (minimale) voorbereiding kan bieden om W&T-onderwijs school-breed te verankeren in de dagelijkse schoolpraktijk. Daarbij merken we op dat zelfs schoolteams die zeer toegewijd zijn om W&T-onderwijs te implementeren, niet noodzakelijkerwijs bij aanvang positieve attitudes hebben ten aanzien van het onderwijzen van W&T. In onze ogen moet attitudetraining moet daarom worden beschouwd als een essentieel en expliciet startpunt van W&T-gerichte professionalisering van basisschoolleraren.

Onze resultaten tonen bovendien dat attitudeverandering en onderwijsvernieuwing niet van de ene dag op de andere dag plaatsvinden en dat daarom schoolontwikkeling op dit gebied substantiële leerkrachtkrachtprofessionalisering vereist (met een tijdsduur van tenminste een heel schooljaar). Hoewel scholen waarschijnlijk de voorkeur geven aan enkele, korte en praktijkgerichte workshops, suggereren onze bevindingen dat dergelijke workshops waarschijnlijk weinig verschil zullen maken in het denken en doen van leraren. Tot slot tonen resultaten het belang van een 'schoolbrede' aanpak voor de implementatie van W&T-onderwijs in basisscholen, omdat een dergelijke benadering samenwerking tussen leraren stimuleerde en de ontwikkeling van gedeelde positieve normen voor onderwijsvernieuwing.

Studie 3 is geaccepteerd voor publicatie als: Post, T., & Walma van der Molen, J. H. (in press). Effects of a longitudinal school development program on primary teachers' attitudes towards inquiry teaching and their inquiry teaching practices. *Teacher College Records.*

Studie 4: De effecten van een schoolontwikkelingsproject gericht op het verbeteren van de attitudes van basisschoolleerlingen ten aanzien van nieuwsgierigheid, zelfbeeld en hun motivaties voor onderzoekend denken

Voortbouwend op de uitkomsten van studie 3, werden vervolgens in studie 4 de effecten onderzocht van het verbeterde lesgedrag van de leraren in de midden- en bovenbouw op de attitudes van hun leerlingen ten aanzien van nieuwsgierigheid, hun impliciete opvattingen over de ontwikkelbaarheid van hun denkvermogen, hun impliciete opvattingen over de invloed van hun inzet op hun leerprestaties, en hun prestatiegerichte motivaties voor onderzoekend en ontwerpend leren. Deze variabelen worden in de literatuur voorgedragen als relevante uitkomstmaten van leerlingontwikkeling op het gebied van W&T in het basisonderwijs.

Hiervoor werden voorafgaand, meteen na afloop en één jaar na het nascholingsprogramma kwantitatieve gegevens van alle groep 6, 7 en 8 leerlingen verzameld met behulp van (Likert-type) vragenlijsten. De effecten van de nascholing op de ontwikkeling van de leerlingen werden onderzocht op basis van een longitudinale, experimentele pretest-posttest onderzoeksopzet met een controlegroep. Daarbij werden Structural Equation Modeling (SEM) technieken ingezet om de relaties tussen de attitudes, opvattingen en motivaties van leerlingen te onderzoeken.

De uitkomsten van de studie tonen dat de attitudes van leerlingen ten aanzien van nieuwsgierigheid, en hun impliciete opvattingen over hun denkvermogen, inderdaad voorspellend lijken voor hun inzet en prestatiegerichte motivaties voor onderzoekend en ontwerpend leren. Daarmee bieden we nieuw inzicht in het psychologische mechanisme dat mogelijk ten grondslag ligt aan het potentiële onderzoeksgedrag van leerlingen in de klas. Het stimuleren van positieve attitudes ten aanzien van nieuwsgierigheid en impliciete overtuigingen over hun leervermogen lijkt daarmee dus een waardevolle aanvulling op bestaande vormen van W&T-gericht onderwijs. De resultaten laten bovendien zien dat de nascholing kleine, maar significante verbeteringen teweegbracht in deze attitudes, opvattingen en motivaties van de leerlingen gedurende de twee jaar van het programma. De bevindingen van studie 4 leveren hiermee aanvullend bewijs voor de effectiviteit van het verrichtte nascholingsprogramma, namelijk ook op het niveau van de leerlingen.

Als we leerlingen dus willen helpen opgroeien tot nieuwsgierige, creatieve en zelfverzekerde denkers, zoals wordt gepromoot door 21[e]-eeuwse onderwijsstandaarden, suggereren de resultaten van deze studie dat leraren leerlingen in de eerste plaats dus bewust moeten maken van *waarom* zij op school nieuwsgierig en onderzoekend moeten zijn. Dit betekent dat leraren tijdens reguliere lessen taal, geschiedenis, wiskunde, en wereldoriëntatie hun leerlingen de educatieve waarde moeten laten inzien van het stellen van onderzoeksvragen. Zij moeten hen duidelijk maken dat hun doorzettingsvermogen om complexe vraagstukken op te lossen hun potentiële denkvermogen vergroot. Daartoe kunnen leraren, zoals aangetoond in deze studie, leerlingen bewuster maken van hun (impliciete) overtuigingen over nieuwsgierigheid en hun denkvermogen en hen vervolgens uitdagen om meer positieve overtuigingen over deze zaken te adopteren, zoals middels reflectie en groepsdiscussie.

Samenvattend, dienen de uitkomsten van deze studie de verdere ontwikkeling van schoolontwikkelingsprojecten gericht op het stimuleren van het onderzoekende denken van leerlingen in het basisonderwijs, namelijk door hun attitudes ten aanzien van nieuwsgierigheid en hun impliciete overtuigingen over hun denkvermogen te bevorderen.

Studie 4 is opgestuurd voor publicatie als: Post, T., & Walma van der Molen, J. H. (2019). Effects of an inquiry-focused school improvement program on the development of pupils' attitudes towards curiosity, their implicit ability and effort beliefs, and goal orientations. Manuscript submitted for publication.

Conclusies

Dit proefschrift toont aan dat een school-brede benadering voor de implementatie van W&T in het basisonderwijs, in combinatie met uitgebreide attitude-gerichte en didactische leerkrachtprofessionalisering, met een bredere focus op W&T-onderwijs, basisscholen kan helpen om W&T-onderwijs te integreren in de eigen schoolpraktijk. Hiermee levert het proefschrift empirisch bewijs voor de geldigheid van de eerdergenoemde hoofdadviezen van de Verkenningscommissie.

Tegelijkertijd onderstrepen onze bevindingen dat het niet eenvoudig is voor basisscholen om W&T en de bijbehorende pedagogiek voor onderzoekend en ontwerpend leren te implementeren. Het veranderen van gewoonten is moeilijk, vooral wanneer zulke verandering van leraren vraagt nieuwe complexe lesvaardigheden te beheersen. Wij denken dat wetenschappelijke vooruitgang op dit gebied dus alleen kan worden bereikt wanneer onderzoekers zich bewust zijn van de complexiteit van W&T-gerichte interventies. Zij moeten expliciete en onderbouwde beslissingen nemen over welke variabelen zij op verschillende schoolniveaus willen bevorderen en meten, en waar mogelijk gebruik maken van getoetste interventies en gevalideerde meetinstrumenten.

Andere belangrijke uitdagingen resteren. Voor het opschalen van W&T-gerichte interventies van kleinschalige veelbelovende voorbeelden naar meer grootschalige onderwijsvernieuwing, moeten grotere kwesties en beleidsmaatregelen worden overwogen. Wanneer bijvoorbeeld de positieve ontwikkeling van de beelden en attitudes van leerlingen ten aanzien van nieuwsgierigheid zou worden opgenomen in nationale onderwijsstandaarden (zoals bepleit wordt in dit proefschrift), zou dit waarschijnlijk een sterke stimulans zijn voor schoolstichtingen en leraren om de ontwikkeling en beoordeling ervan te benadrukken als een integraal onderdeel van hun reguliere schoolpraktijk. Overheidsinstanties en onderwijsorganisaties zouden daarom meer onderzoek moeten subsidiëren dat gericht is op het beter definiëren en ontwikkelen van beoordelingsmaten van dergelijke leerlingkwaliteiten, aanvankelijk voor onderzoeksdoeleinden, en later voor formatieve en summatieve beoordeling (zie ook Geisinger, 2016; Inspectie van het Onderwijs, 2019).

Alleen als het stimuleren van het onderzoekende en ontwerpende denken van leerlingen door basisschoolleerkrachten wordt gezien als een integraal onderdeel van hun lespraktijk, niet minder dan lesgeven over het oplossen van een wiskundig probleem of het schrijven van een werkstuk, zullen we waarschijnlijk pas wezenlijke verandering zien in hun dagelijkse lesgeven. Leraren moeten eerst zélf overtuigd

raken van het nut, het plezier en de mogelijkheid om nieuwe dingen te leren door te onderzoeken en te ontwerpen. Daardoor zijn in staat zijn als authentieke rolmodellen te fungeren voor hun leerlingen. Zij kunnen dan de perceptie bij leerlingen cultiveren dat school niet alleen draait om het onthouden van feiten, het verzamelen van informatie op het internet, of het berekenen van oppervlakten. Maar dat school uiteindelijk gaat over het worden van een nieuwsgierige ontdekkingsreiziger, een originele denker en een veerkrachtige concurrent op welk gebied zij zich in de toekomst ook zullen bevinden. Leraren zullen leerlingen moeten uitleggen waarom hun leraren hen tijdens hun hele academische carrière zullen blijven uitdagen met nieuwe en complexe lesinhouden, wat zijzelf moeten investeren om te excelleren in hun leven en wat school hen brengt als zij zich werkelijk committeren.

Op deze manier kan school een plek worden waar het nieuwsgierige, creatieve en gedurfde denken van leerlingen niet alleen overleeft, maar floreert.

Contributions

Contributions

International refereed journals

Post, T., & Walma van der Molen, J. H. (2018). Do children express curiosity at school? Exploring children's experiences of curiosity inside and outside the school context. *Learning, Culture and Social Interaction, 18*, 60–71.

Post, T., & Walma van der Molen, J. H. (2018). Development and validation of a questionnaire to measure primary school children's images of and attitudes towards curiosity (the CIAC questionnaire). *Motivation and Emotion, 43*, 159–178. This paper received the PhD Best Paper Award 2017 by the Motivation & Emotion special interest group of the European Association for Research on Learning and Instruction (EARLI) conference in Finland.

Post, T., & Walma van der Molen, J. H. (in press). Effects of a longitudinal school development program on primary teachers' attitudes towards inquiry teaching and their inquiry teaching practices. *Teacher College Records.*

Post, T., & Walma van der Molen, J. H. (2019). Effects of an inquiry-focused school improvement program on the development of pupils' attitudes towards curiosity, their implicit ability and effort beliefs, and goal orientations. Manuscript submitted for publication.

Conference presentations

Post, T., & Walma van der Molen, J. H. (2017). *Measuring Primary Children's Images of and Attitudes Towards Curiosity.* Paper presentatie op de European Association for Research on Learning and Instruction (EARLI) conferentie, Tampere, Finland, 2 augustus. Dit paper ontving de PhD Best Paper Award van de Motivation & Emotion special interest group.

Post, T., & Walma van der Molen, J. H. (2015). *Creating innovative and powerful learners: A multidisciplinary framework.* Paper presentatie op de conferentie European Association for Research on Learning and Instruction (EARLI), Limassol, Cyprus, 29 augustus.

Post, T., & Walma van der Molen, J. H. (2015). *Naar een cognitief psychologisch raamwerk voor 21e-eeuws leren.* Poster presentatie tijdens de Onderwijs Research Dagen (ORD), Leiden, Nederland, 19 juni.

Post, T. (2016). *Towards an interdisciplinary framework of 21st-century learning ability.* JURE, Helsinki, Finland, 8 juli.

Workshops and invited lectures

Post, T. (2019). *Nieuwsgierigheid als motor van leren.* Lezing op het congres De Staat van het Onderwijs, Utrecht, Nederland, 10 april.

Post, T. (2019). *Onderwijs voor morgen.* Lezing op de onderwijsdag van scholenstichting Archipel, Zutphen, Nederland, 27 maart.

Post, T. (2019). *Onderwijs voor morgen.* Lezing op de netwerkbijeenkomst van Kenniscentrum Wetenschap & Technologie Oost, Zutphen, Nederland, 6 maart.

Post, T. (2018). *Onderwijs voor morgen.* Lezing voor de onderwijsdag van scholenstichting Consent, Enschede, Nederland, 20 november.

Post, T. (2018). *Onderzoekende en ontwerpende schoolteams.* Lezing voor de masteropleiding Educational Leadership (MEL) van Saxion Hogeschool, Enschede, Nederland, 8 november.

Post, T. (2018). *Onderzoeken en ontwerpen met muziek.* Lezing voor het Meer Muziek in de Klas symposium, Ede, Nederland, 1 november.

Post, T. (2018). *Belang van bedrijfsbezoeken en voorwaarden.* Lezing op een netwerkbijeenkomst van Kenniscentrum Wetenschap & Technologie Oost (KWTO), Almelo, Nederland, 27 september.

Post, T. (2018). *Belang van bedrijfsbezoeken en voorwaarden.* Lezing voor de Week van de Techniek, Almelo, Nederland, 20 september.

Post, T. (2018). *Onderwijs voor morgen.* Lezing voor Cultuurmij Oost, Burgers Zoo, Nederland, 29 april.

Post, T. (2018). *Nieuwsgierigheid als motor voor leren.* Twents Meesterschap conferentie, Nederland, Enschede, 24 januari.

Post, T. (2017). *Talentontwikkeling in het onderwijs.* Techniekpact conferentie, Hengelo, Nederland, 16 november.

Post, T. (2017). *Effectieve schoolexcursies naar technische bedrijven.* Keynote op de onderwijsdag van scholenstichting Op Kop, Steenwijk, Nederland, 18 oktober.

Post, T. (2016). *Innovatieve en nieuwsgierige leerlingen voor morgen.* Workshop tijdens onderwijsdag Stichting KOE, Enschede, Nederland, 15 maart.

Post, T. (2016). *Wat kunnen leerkrachten doen om het beste uit hun leerlingen te halen?* Workshop op de conferentie Twents Meesterschap, Enschede, Nederland 28 januari.

Post, T. (2015*). Randvoorwaarden voor cultuureducatie.* Paneldiscussie bij de De Cultuur Loper conferentie in het Provinciehuis, 's Hertogenbosch, Nederland, 7 oktober.

Post, T. (2015). *Innovatieve leerlingen voor morgen.* Keynote op Iselinge Hogeschool, Doetinchem, Nederland, 3 juni.

Walma van der Molen, J. H., Van Aalderen-Smeets, S. I., & Post, T. (2015). *Research center for Science Education and Talent Development (SETD): Stimulating and tracking children's talent development.* Poster presentatie op het congres Learning and the Brain, New York, Verenigde Staten, 8 mei.

Post, T. (2014). *Coachende ouders: Kinderen boeien en uitdagen met technologie.* Lezing op een symposium van Kenniscentrum Wetenschap & Technologie Oost, Epe, Nederland, 19 november.

ICO

ICO Dissertation Series

In the ICO Dissertation Series dissertations are published of graduate students from faculties and institutes on educational research within the ICO Partner Universities: Eindhoven University of Technology, Leiden University, Maastricht University, Open University of The Netherlands, University of Amsterdam, University of Twente, Utrecht University, VU University Amsterdam, and Wageningen University, and formerly University of Groningen (until 2006), Radboud University Nijmegen (until 2004), and Tilburg University (until 2002). The University of Groningen, University of Antwerp, University of Ghent, and the Erasmus University Rotterdam have been 'ICO 'Network partner' in 2010 and 2011. From 2012 onwards, these ICO Network partners are full ICO partners, and from that period their dissertations will be added to this dissertation series.

375. Okkinga, M. (02-02-2018) Teaching reading strategies in classrooms – does it work? Enschede: University of Twente.

376. Thomsen, M. (09-02-2018) Teachers Trust. Measurement, sources and consequences of teacher's interpersonal trust within schools for vocational education and training. Amsterdam: University of Amsterdam.

377. Van der Wurff, I.S.M. (09-02-2018) Fatty acids, Cognition, School Performance and Mental Well-Being in Children and Adolescents. Heerlen: Open University of The Netherlands.

378. Raaijmakers, S.F. (16-02-2018) Improving self-regulated learning: Effects of training and feedback on self-assessment and task-selection accuracy. Utrecht: Utrecht University.

379. Zhao, X. (07-03-2018) Classroom assessment in Chinese primary school mathematics education. Utrecht: Utrecht University.

380. Van Rooij, E.C.M. (15-03-2018) Secondary school students' university readiness and their transition to university. Groningen: University of Groningen.

381. Vanlommel, K. (26-03-2018) Opening the black box of teacher judgement: the interplay of rational and intuitive processes. Antwerp: University of Antwerp.

382. Boevé, A.J. (14-05-2018), Implementing Assessment Innovations in Higher Education. Groningen: University of Groningen.

383. Wijsman, L.A. (30-05-2018) Enhancing Performance and Motivation in Lower Secondary Education. Leiden: Leiden University.

384. Vereijken, M.W.C. (22-05-2018) Student engagement in research in medical education. Leiden: Leiden University.

385. Stollman, S.H.M. (23-05-2018) Differentiated instruction in practice: A teacher perspective. Leiden: Leiden University.

386. Faddar, J. (11-06-2018) School self-evaluation: self-perception or self-deception? Studies on the validity of school self-evaluation results. Antwerp: University of Antwerp.

387. Geeraerts, K. (25-06-2018) Dood hout of onaangeboorde expertise? Intergenerationale kennisstromen in schoolteams. Antwerp: University of Antwerp.

388. Day I.N.Z. (28-06-2018), Intermediate assessment in higher education. Leiden: Leiden University.

389. Huisman, B.A. (12-09-2018) Peer feedback on academic writing. Leiden: Leiden University.

390. Van Berg, M. (17-09-2018) Classroom Formative Assessment. A quest for a practice that enhances students' mathematics performance. Groningen: University of Groningen.

391. Tran, T.T.Q. (19-09-2018) Cultural differences in Vietnam: differences in work-related values between Western and Vietnamese culture and cultural awareness at higher education. Leiden: Leiden University.

392. Boelens, R. (27-09-2018) Studying blended learning designs for hands-on adult learners. Ghent: Ghent University.

393. Van Laer, S. (4-10-2018) Supporting learners in control: investigating self-regulated learning in blended learning environments. Leuven: KU Leuven.

394. Van der Wilt, F.M. (08-10-18) Being rejected. Amsterdam: Vrije Universiteit Amsterdam.

395. Van Riesen, S.A.N. (26-10-2018) Inquiring the effect of the experiment design tool: whose boat does it float? Enschede: University of Twente.

396. Walhout, J.H. (26-10-2018) Learning to organize digital information Heerlen: Open University of The Netherlands.

397. Gresnigt, R. (08-11-2018) Integrated curricula: An approach to strengthen Science & Technology in primary education. Eindhoven: Eindhoven University of Technology.

398. De Vetten, A.J. (21-11-2018) From sample to population. Amsterdam: Vrije Universiteit Amsterdam.

399. Nederhand M.L. (22-11-2018) Improving Calibration Accuracy Through Performance Feedback. Rotterdam: Erasmus University Rotterdam.

400. Kippers, W.B. (28-11-2018) Formative data use in schools. Unraveling the process. Enschede: University of Twente.

401. Fix, G.M. (20-12-2018) The football stadium as classroom. Exploring a program for at-risk students in secondary vocational education. Enschede: University of Twente.

402. Gast, I. (13-12-2018) Team-Based Professional Development – Possibilities and challenges of collaborative curriculum design in higher education. Enschede: University of Twente.

Dankwoord

Dankwoord (acknowledgments)

Ineens besef ik me dat dit het laatste stuk is dat ik voor mijn proefschrift schrijf. De afsluiting van een bijzondere periode in mijn leven. Een periode die ik heb beleefd als een inspirerende en wervelende vlucht door de wereld van de wetenschap. Een spannende vlucht langs allerlei toppen en dalen. Soms met uitzicht op zonnige stranden, soms met uitzicht op donderswolken. Soms verdwaalde ik, maar gelukkig vond ik de juiste weg terug naar mijn bestemming. Sommigen van jullie zijn met me meegereisd, terwijl anderen mijn reis vanaf de grond hebben gevolgd en aangemoedigd. Zonder jullie had ik mijn bestemming nooit bereikt.

Prof. dr. Walma van der Molen, beste Juliette, mijn promotor en dagelijkse begeleider, het eerste en grootste dankwoord aan jou. Onze besprekingen zorgden altijd dat zaken weer op scherp werden gezet. Je hield me op koers, voorzag me in de nodige landkaarten en wist me uit de wolken te leiden wanneer ik verdwaalde. Daarbij heb ik je kritische denken, nieuwsgierigheid en aanmoediging enorm gewaardeerd. In het bijzonder herinner ik me onze ontwikkeling en validatie van de CIAC-vragenlijst, en de prijs die we voor dit onderzoek ontvingen in Finland (wat was die studie interessant!). Ik voel me twee keer zo sterk als bij aanvang van dit project. Ook bedank ik je voor de kans en stimulans die je me gaf voor het verzorgen van diverse lezingen, workshops en andere bijdragen voor het Nederlandse onderwijsveld. Daarop bouw ik de komende jaren voort.

Ook bedank ik graag mijn promotiecommissie, bestaande uit prof. dr. De Jong, prof. dr. Meijer, prof. dr. De Vries, prof. dr. Bohlmeijer en dr. Steenbeek voor het beoordelen van dit proefschrift en hun deelname aan de promotieplechtigheid.

In het bijzonder bedank ik ook Jos Brunninkhuis en Maria Hendriks als directie van TechYourFuture, het expertisecentrum voor techniekonderwijs. TechYourFuture blies leven in dit promotieonderzoek. Bedankt voor jullie financiële steun, vertrouwen en advies.

Graag bedank ik ook alle schoolteams die hebben deelgenomen aan het onderzoek en, in het bijzonder, alle schoolleiders, leraren en leerlingen die ik in de verschillende studies heb mogen bevragen. Zonder jullie vertrouwen en openhartigheid had ik de effecten en complexiteit van W&T-gerichte schoolontwikkeling niet kunnen onderzoeken. Daarmee hebben jullie een essentiële rol gespeeld in de realisatie van dit promotieonderzoek en de ontwikkeling van belangrijke wetenschappelijke kennis. Daarvoor bedank ik jullie nadrukkelijk.

Veel dank ook aan Erik Groot Koerkamp voor de vruchtbare samenwerking en Gerard Venneman en Patty van Scherpenzeel van het KWTO voor de regionale kennisverspreiding tijdens het promotietraject. Ik hoop nog veel met jullie te mogen werken in de komende jaren.

Collega's van vakgroep ELAN, het was fijn omringd te zijn door zulke enthousiaste en bevlogen mensen tijdens mijn reis. In het bijzonder bedank ik mijn (vroegere) kamergenoten Sandra, Frances, Niels en Kim. Bedankt voor de bizarre bedenksels, handstanden, propjes, slechte grappen en lunchwandelingen. Sandra en Frances, mijn paranimfen, bedankt voor jullie aanmoediging en steun. 'Uiteindelijk draait alles om …'

Dennis en Hannah, ook bedank ik jullie voor de aanmoediging tijdens mijn promotietraject. Dennis, bedankt voor het grenzeloze mentale 'sparren' over alles wat belangrijk is en de jarenlange vriendschap die we inmiddels hebben opgebouwd sinds de start van mijn opleiding aan de Universiteit Twente. Ik hoop nog heel lang van je hersenspinsels en nieuwsgierigheid te mogen genieten.

Pap en mam, Anneloes en Eline, bedankt voor jullie vertrouwen en aanmoediging. Jullie waren voor mij vaak voorbeelden bij het bestuderen van de nieuwsgierigheid, creativiteit en verbeeldingskracht van kinderen. Jullie hebben me een hoop wijsheid en ervaring op dit vlak vanuit huis meegegeven. Evengoed bedank ik mijn schoonfamilie voor alle steun en aanmoediging. Ik hou van jullie allemaal.

Mijn vriendin Mireille, ik kan je niet genoeg bedanken. Je was er altijd voor me om mijn ideeën, frustraties en vragen aan te horen en hierover met me te sparren. Je liefde maakt me elke dag sterker en gelukkiger. Je geeft 'schwung' aan mijn leven. Ik hoop dat we samen nog heel veel mooie momenten mogen meemaken, zoals deze. Ik hou heel veel van je.

Tim Post