## Teaching towards historical expertise

*Developing students' ability to reason causally in history*

Stoel, G.L.

**Publication date**
2017
**Document Version**
Other version
**License**
Other

[Link to publication](#)

# CHAPTER 6

## SUMMARY, GENERAL CONCLUSIONS AND DISCUSSION

In this final chapter, we will first summarize the five chapters in this dissertation. Subsequently we will connect the main findings from the four studies and draw some conclusions about the effectiveness of explicit teaching, and about the additional demands of writing a historical explanation. Furthermore, we will reflect on the question how the MDL and the framework of historical reasoning supported and appended each other in our operationalization of teaching and learning causal historical reasoning. In the final paragraphs, we will discuss methodological considerations, make some suggestions for future research, and present the implication for practice that we draw from our studies.

## SUMMARY

### CHAPTER 1. GENERAL INTRODUCTION

Chapter 1 introduced the background and goals of the studies in this dissertation. We contended that in current history curricula, students are not only expected to acquire historical knowledge, but also develop the critical skills to assess and construct historical accounts (Erdmann & Hassberg, 2011). The ability to answer causal historical questions is one of these competencies (van Drie & van Boxtel, 2008). However, in many history classrooms, teaching and learning activities remain primarily focused on the acquisition and reproduction of first-order knowledge and historical reasoning skills often remain implicit (van Boxtel, van Drie, & Kropman, 2010; VanSledright, 2011).

The studies in this dissertation were designed to contribute to bridging this gap between the intended curriculum and educational practice. Towards this

goal, our first aim was to define a pedagogical approach that we expected would foster 11th grade preuniversity students' ability to construct a causal historical explanation. To define this approach, we used the model of domain learning (*MDL*; Alexander, 2003, 2005) and combined this model with the framework of historical reasoning (van Drie & van Boxtel, 2008). We expected the MDL to provide a suitable framework because, contrary to other models that primarily focus on defining pedagogical principles for teaching (domain-specific) higher-order reasoning, the MDL was developed as a conceptualization of the cognitive and affective dimensions involved in domain-specific reasoning and of the development of expertise. This model matched, and appended, the framework of historical reasoning that focused primarily on defining the domain-specific components involved in answering a historical question (cf. van Drie & van Boxtel, 2008). Only in second instance, did Alexander, based on the MDL, define instructional principles of learning environments designed to foster expertise (2005). Consequently, the model also allowed us to define a pedagogical approach connected to our domain-specific learning goals. No previous studies in history education had used the MDL to operationalize teaching and learning.

Important background principles in our pedagogical approach were—working with realistic open tasks, interchanging group work and classroom discussion, making thinking visible, and raising situational interest. However, in this dissertation we focused specifically on the role of explicit teaching as an indispensable principle to foster causal historical reasoning. Although several studies in history education had shown the effectiveness of explicit teaching on students' ability to work with historical sources (e.g. De La Paz, 2005; Nokes et al., 2007; Reisman, 2012), no previous studies had focused on the explicit teaching of causal historical reasoning. Based on the MDL, we extended the focus of explicit teaching to also include explicit teaching of second-order concepts (e.g. causal categories and vocabulary) and explicit reflection on epistemological questions (e.g. about the interpretative nature of causal explanations and

methodological criteria for evaluating these claims). The research-questions in this dissertation centered around the effectiveness of explicit teaching (along these dimensions) on students' ability to reason causally in history.

The questions in this dissertation were as follows:

1.  What constitutes causal historical reasoning and what are the effects of our design-principles—derived from the model of domain learning—in general and of explicit instruction in particular, on the causal historical reasoning ability of 11th grade preuniversity students? (*Chapter 2*)

2.  What are the effects of explicit teaching of second-order concepts, causal reasoning strategies, and epistemological underpinnings (in the context of a collaborative explanatory task) on 11th grade students' (a) second order and strategy knowledge, (b) their epistemological beliefs and (c) their ability to construct a causal explanation? In addition: what are the effects on first-order knowledge and individual interest? (*Chapter 3*)

3.  Which aspects of causal historical reasoning do students include in their essay revisions after a lesson-unit in which second-order concepts and strategies related to historical causation, as well as epistemological reflection on the nature of historical explanations, are explicitly taught? And how are these revisions related to students' initial text-structure? (*Chapter 4*)

4.  Does our epistemological beliefs questionnaire on historical knowledge and knowing confirm a theoretical model in which epistemological beliefs in history are primarily divided between (a) naïve epistemological beliefs and (b) nuanced epistemological beliefs? (*Chapter 5*)

In the first study, we defined a domain-specific instructional framework in which we operationalized causal historical reasoning and the pedagogical principles, and designed a lesson-unit on World War I for 11th grade preuniversity students. In secondary school, World War I is a classic subject to teach about causality.

Subsequently, we conducted a quasi-experimental study with an explicit and an implicit teaching condition, to explore the effects of our principles on students' ability to reason causally in history. The focus of our analysis lay in the added effects of explicit instruction. In the second study, we replicated the intervention as a randomized controlled experiment with an experimental and a control group. In this study, our research question focused specifically on the effects of explicit instruction and two open prompts were added to explore students' reflections on learning gains. In the third study, we analyzed in detail how students revised their essays after the lesson-unit and investigated which aspects of causal historical reasoning they were able to include in their texts. In the final study, we developed and tested an alternative questionnaire for measuring students' epistemological beliefs in history.

CHAPTER 2. TEACHING CAUSAL REASONING IN HISTORY.

*RQ: What constitutes causal historical reasoning and what are the effects of our design-principles—derived from the model of domain learning—in general and of explicit instruction in particular, on the causal historical reasoning ability of 11th grade preuniversity students?*

*Theoretical background*

Two theoretical models underpinned our definition of causal historical reasoning—the model of domain learning (MDL) and the domain-specific framework of historical reasoning. Based on these models, we distinguished between historical (or first-order) knowledge, knowledge of second-order concepts (VanSledright & Limón, 2006), (domain-specific) deep-level strategies, and epistemological beliefs. We defined causal historical reasoning as—"an activity in which a person *constructs a historical explanation* by using *first-order knowledge,* and *knowledge of second-order concepts* and *strategies* related to causality,

*provides arguments* and *counterarguments* to support causal statements, and uses *evidence* taken from *historical sources*. The approach and outcome of this activity is influenced by the *epistemological beliefs* about the nature of historical causation a student holds and his *interest* in history."

Based on literature, we operationalized the concepts, strategies and epistemological beliefs involved in causal historical reasoning. Examples of important strategies for constructing historical explanations are—looking for multiple causes, modeling causal relations in non-linear and complex ways, and categorizing causes along multiple dimensions, such as time, domain, and role (Chapman, 2003; Coffin, 2004; Halldén, 1997; Seixas & Morton, 2014). In these causal categories, second-order concepts play an important role. Examples of these concepts are—"long-term or short-term", "direct or indirect" (related to time); "economic", "social", "political" (related to domain); and "trigger", "catalyst", and "precondition" (related to the role causes play). An important epistemological characteristic of causal historical reasoning that students often fail to acknowledge is the fact that causes and causal relations are interpretative constructions based on disciplinary criteria of evidence and argument (Lee & Shemilt, 2009; Maggioni et al., 2009).

Qualitative research on history education has described three epistemological positions that students can hold on the nature of historical knowledge and knowing (e.g. Lee & Shemilt, 2003; Maggioni et al, 2009). In the 'copier' (or objectivist) stance, students regard historical knowledge as fixed and hold objective truth about the past to be possible. In the 'borrower' (or subjectivist) stance, students understand that fixed knowledge about the past is not possible, because the past is gone. However, these students lack an understanding of the methodological criteria to assess the reliability of historical claims. Therewith reducing historical interpretations to mere 'opinions'. The most nuanced position is the 'criterialist' stance. In this stance, students

understand the interpretative nature of historical knowledge, but also understand the domain-specific criteria for assessing the reliability of historical claims.

In a related article, Alexander also outlined general characteristics of a learning environment designed to foster domain-specific expertise. We elaborated these characteristics, by combining them with other models of teaching and learning (e.g. Collins et al., 1991; Merill, 2002) and with research on history education. Five pedagogical principles were defined—(a) work on open-ended realistic tasks (cf. van Drie et al., 2006); (b) organize social interaction (cf. Pontecorvo & Girardet, 1993; Del Favero et al., 2007); (c) make thinking visible by working with graphical organizers and concept maps (cf. Prangsma et al., 2009; van Drie, et al., 2005); and (d) raise situational interest or rooted relevance (cf. Barton & Levstik, 2004; Del Favero et al., 2007). Finally, we discerned explicit teaching of domain-specific strategies and second-order concepts, as well as explicit reflection on epistemological questions as a key principle in our learning environment (cf. De La Paz, 2005; Khishfe & Abd-El-Khalick, 2002).

*Designing the lesson-unit*

Based on these characteristics, we designed a lesson-unit (three lessons) in which we 'inked in' our design principles with concrete teaching and learning activities, partly derived from practitioner-oriented literature. We designed the lesson-unit in two conditions—an *explicit* experimental condition and an *implicit* control condition. The central question in both conditions was "how can we explain the outbreak of World War I?" In both conditions, students worked in triads on (sub)tasks to answer the inquiry question, and card sorting and the construction of graphical representations were central activities. In both conditions, students presented and discussed their explanations in a whole-class setting. The main difference between the conditions was that students in the implicit condition worked on the inquiry question while feedback and instruction primarily focused on first-order knowledge. In contrast, in the explicit condition explicit feedback

and instruction was given on the underlying strategies, concepts, and epistemological ideas connected to causal historical reasoning.

*Developing the research instruments*

In line with the MDL, we took a multi-dimensional approach to assessing students' improvement in causal historical reasoning. (a) We developed a test to assess students' first-order knowledge about events, concepts, and chronology before World War I. (b) We designed a fictional story to test students' ability to apply second-order concepts. (c) We developed a questionnaire assessing knowledge of causal historical strategies. (d) Students' epistemological beliefs about history were measured with the Beliefs about History Questionnaire (BHQ; Maggioni et al., 2004, 2009). Finally, (e) we designed two document-based essay-tasks that asked students to write a 300-word explanation on the question how (i) Germany or (ii) Russia became involved in World War I. The essay tasks intended to measure students' ability to apply causal historical reasoning.

*Results of the study*

In a quasi-experimental study, three 11th grade preuniversity history classes from one school were divided over an *explicit* condition (two classes; $N = 50$) and an *implicit* condition (one class; $N = 24$). In line with our expectations, results showed that students in both conditions acquired first-order knowledge and that this acquisition did not differ between conditions. Furthermore, as we had hypothesized, students in the explicit condition acquired significantly more knowledge of second-order concepts and causal strategies. Explicit attention and practice appeared to be a precondition for developing this knowledge. The study did not show clear effects on students' epistemological beliefs. Statistical analysis of the BHQ yielded an unacceptable low reliability for the factor measuring criterialist ideas. A second – reliable – factor combined items measuring

objectivist and subjectivist beliefs and did not show any changes between pre- and post-test. Contrary to our expectations, we found no differences between conditions on the writing task, although students' explanations in both condition significantly improved in the post-test.

The study led us to conclude that the MDL does provide an effective framework for conceptualizing causal historical reasoning and defining pedagogical principles, as well as for assessing development of students' causal reasoning. However, we also concluded that applying the knowledge of causal concepts and strategies in a document-based writing task was difficult for students. We hypothesized that a writing task might require additional knowledge and skills (e.g. genre-knowledge and knowledge of working with sources) that students did not possess or could not (yet) integrate with their enhanced understanding of causal historical reasoning. This might require more practice and the explicit teaching of additional components of historical reasoning and of writing historical texts.

CHAPTER 3. EFFECTS OF EXPLICIT TEACHING ON STUDENTS' CAUSAL REASONING

*RQ: What are the effects of explicit teaching of second-order concepts, causal reasoning strategies, and epistemological underpinnings (in the context of a collaborative explanatory task) on 11th grade students' (a) second order and strategy knowledge, (b) their epistemological beliefs and (c) their ability to construct a causal explanation? In addition: what are the effects on first-order knowledge and individual interest?*

The study described in chapter 3, replicated the first intervention study, but focused specifically on the effects of explicit teaching. An important difference with the first study was that we designed this study as a randomized controlled trial. We limited the amount of learning goals and strove to further increase the

explicitness of the learning activities. Because of the importance of interest in the MDL, we added a questionnaire measuring individual interest in history (based on Linnenbrink-Garcia et al., 2010; Pintrich et al., 1993). Based on the first study, we used only two of the three scales of the BHQ to measure epistemological beliefs (namely, the subjectivist and criterialist scales; Maggioni et al., 2004, 2009). Further, we added two open prompts to the post-test that asked students to reflect on their learning gains, and describe a heuristic for answering causal questions in future history lessons. Finally, at post-test we asked students to revise their pre-test essay about the involvement of Germany in the war. The choice for revision, instead of writing a new essay, was made to reduce the complexity of the writing task and heighten the sensitivity of the instrument.

95 11th–grade preuniversity students (four classes from one school) took part in the experiment. Students were randomly divided over a treatment ($N_{exp}$ = 53) and a control condition ($N_{contr}$ = 42). Results showed that students in the experimental group scored significantly higher at the posttest on knowledge of causal-reasoning strategies and second-order concepts. Furthermore, the experimental condition attributed a significantly higher value to both epistemological scales (subjectivist and criterialist). The increased score on subjectivism contrasted with our expectation. Third, students' answers on the open prompts were significantly more domain-specific in the experimental group. On the heuristic-prompt, 58% of these students referred to at least one aspect of causal historical reasoning, versus 21% in the control group. This same difference was found for reflections on learning gains. An interesting finding in these reflections was that 23% of the experimental group referred to epistemological aspects of the lessons (versus 2% in the control group).

Just like in the first intervention study, and contrary to our hypothesis, we found no differences between conditions in the overall quality of students' written explanations. However, the experimental group did score significantly higher on one core criterion of the scoring-rubric—"use of second-order

language and causal connections". As expected, this study found no between-group differences on (the development of) first-order knowledge. Finally, students in the experimental condition reported a higher amount of individual interest at post-test than the control group. This was unexpected, because individual interest is usually conceptualized as a relatively stable construct.

An interesting finding in this study was the increased correlation in the post-test of the experimental condition between criterialist epistemological beliefs and individual interest. In the pre-test, as well as in the post-test of the control condition, the correlation between these constructs was moderate. However, in the experimental group the correlation at post-test was (very) strong. This result appeared to be mainly caused by the heightened level of individual interest that students in the explicit condition reported at post-test.

The results of the second study were in line with the first study and deepened our understanding through the addition of more qualitative elements (the open prompts), the individual interest questionnaire and the correlational analysis. We concluded that explicit teaching of causal strategies and second-order concepts is an indispensable precondition for developing causal historical reasoning. Students' answers to the open prompts, as well as the use of second-order concepts in the essays, showed that this knowledge became more concrete in the experimental group. Furthermore, the answers on the open prompts, the increase on individual interest and the correlation between students' interest and their criterialist epistemological beliefs, all indicated that explicating this knowledge and reflecting on the epistemological underpinnings of causal historical explanations contributed to the value that students allotted to the lessons.

CHAPTER 4. AN EXPLORATION OF STUDENTS' EXPLANATORY WRITING IN HISTORY

*RQ: Which aspects of causal historical reasoning do students include in their essay revisions after a lesson-unit in which second-order concepts and strategies related to historical causation, as well as epistemological reflection on the nature of historical explanations, are explicitly taught? And how are these revisions related to students initial text-structure?*

Writing historical explanations based on multiple sources is a highly complex task. Not only do students have to apply strategies and concepts related to causal historical reasoning, but also do they have to combine this with first-order knowledge, knowledge of causal historical text-structures and rhetorical demands (Coffin, 2004; Schleppegrell et al., 2008; van Drie et al., 2014) and with a critical handling of historical sources (Wineburg, 1991). Finally, epistemological beliefs might influence the way in which students perceive their task and the types of strategies they engage in (Buehl & Alexander, 2001; Lee & Shemilt, 2009).

In line with this complexity, both intervention studies revealed limited effects on the quality of students' explanatory essays. This result raised questions about the complexities of causal historical writing and about the exact nature of students' revisions after the lesson-unit. In this qualitative study, we analyzed in detail the types of revisions that students from the experimental group integrated in their original essays. In line with our definition of historical causation, we looked at the second-order concepts and causal strategies that had been explicitly taught, as well as at revisions related to the argumentative quality of the texts—e.g. insertion of claims and conclusions, and reference to sources. Besides, we explored how initial text-structure interacted with the types of revisions students made. Towards this aim, we defined two causal historical text-structures—linear-chronological and nonlinear-thematic (based on Coffin, 2004).

We selected a matched sample of forty essays from the experimental condition (the pre- and post-tests of twenty students). Through open and axial coding, we categorized and counted the revisions related to historical causation. Three main categories were defined: *causal elaboration, second-order concepts,* and *argumentation.* Each category was divided in several subcategories. Revised passages could receive multiple codes. We did not only categorize the revisions, but we also analyzed the domain-specific quality of the changes students made.

Results showed that students (a) integrated *descriptive* and *analytical* causal second-order concepts, (b) elaborated the causal structure of their texts by inserting abstract historical phenomena (*conceptual elaborations*), or concrete historical events (*concrete elaborations*), and (c) inserted argumentative elements that strengthened author-presence in the texts. However, almost all argumentative revisions were coded as *causal claim*. No argumentative revisions were found in which students referred to evidence from sources.

Our analysis also showed that initial text-structure (linear-chronologic or nonlinear-thematic) was related to the amount and type of revisions made. Linear essays on average contained 14 revisions, whereas non-linear essays on average contained 23 revisions. Students who organized their text more thematically included more analytical second-order concepts and made more causal claims. Both text-structures integrated roughly the same amount of causal elaborations. However, we did find a qualitative difference in the 'function' of *conceptual elaborations* (revisions in which students integrated historical phenomena such as *nationalism*). In the thematic essays, these elaborations often transformed the nature of the paragraph by developing the abstract concept as the central theme and 'reducing' concrete events to a more exemplary role. This process of "nominalization" (e.g. Coffin, 2004) was witnessed less often in linear-chronological texts.

Finally, results showed that although many revisions (in both text-structures) were coded as the insertion of a causal claim, these claims tended to

lack argumentative support. This lack of argumentative support was also found for revisions where students added interpretative elements to their text, such as analytic second-order language (e.g. words like "a background cause", or a "trigger"). In the entire sample, only very few warrants for these interpretative elements were coded.

We concluded that students were able to integrate causal concepts and to strengthen their causal explanations by including concrete events and abstract phenomena. Furthermore, in their revisions students strengthened the argumentative quality of their text by inserting causal claims. Simultaneously, our analysis also revealed that other dimensions impacted the quality of students' revisions. First of all, we found that the type and number of revisions students made was related to the initial text-structure (linear versus non-linear) of their essays. Furthermore, students' claims and interpretative revisions lacked argumentative support and almost no references to evidence were made. In their essays, students appeared to primarily use historical sources to 'copy' information to their essays. These findings were in line with earlier studies and might be related to students' beliefs about historical knowledge (Wineburg, 1991; McCarthy Young & Leinhardt, 1998). We contended that knowledge of second-order concepts and strategies is an important aspect to enhance causal writing, but knowledge of causal text-structures, strategies for working with historical evidence, and providing arguments and counter-arguments, should also be explicitly taught if we want students to progress in their ability to write nuanced causal explanations.

CHAPTER 5. MEASURING EPISTEMOLOGICAL BELIEFS IN HISTORY EDUCATION

*RQ: Does our epistemological beliefs questionnaire on historical knowledge and knowing confirm a theoretical model in which epistemological beliefs in history are primarily divided between (a) naïve epistemological beliefs and (b) nuanced epistemological beliefs?*

In chapter 5, we dove deeper into the complexities of measuring epistemological beliefs in history. Both intervention studies had revealed that assessing epistemological beliefs was a complex challenge and that the Beliefs about History Questionnaire (BHQ; Maggioni et al, 2004, 2009) yielded several problems related to the reliability and stability of the scales. In our first study objectivist and subjectivist items were found to load together. In our second study, we witnessed that development on the subjectivist scale was contrary to theory.

This study was designed to develop and test an alternative questionnaire for measuring epistemological beliefs in history. We developed the questionnaire on a theoretical model in which epistemological beliefs were primarily divided between naïve and nuanced beliefs. We related naïve ideas to understanding historical knowledge as a fixed representation of the past and to the belief that historical accounts can only be written if sources are unambiguous and objective (naïve–objective). Furthermore, we related naïve ideas to understanding historical knowledge as mere opinion (naïve–subjective). We related nuanced epistemological ideas about history to understanding historical knowledge as interpretative and temporary (nuanced–subjective) and to valuing methodological criteria for generating reliable historical knowledge (nuanced–criteria). A new aspect of our model was the separation of subjectivist items between naïve and nuanced ideas.

We developed a 26-item questionnaire, scored on a 6-point Likert-scale. The naïve scale contained 15 items and the nuanced scale contained 11 items. Participants in this study were 922 exam students enrolled in their final year of higher general continued education (HG) and preuniversity education (PU). We used exploratory factor analysis (EFA) to analyze the data and tested the stability of the items by comparing the factors with the scores of seven expert historians and philosophers of history. We also compared the differences in mean scores between the two school-levels. Finally, we investigated whether the correlations between criterialist ideas, interest, and first-order knowledge, found in chapter 3 could also be witnessed in this dataset.

Results showed that a two-factor solution oversimplified the underlying structure. EFA extracted five factors that could, however, be interpreted within our theoretical assumptions. On the nuanced side, EFA extracted a factor related to items focusing on methodological criteria for generating historical knowledge. Experts consistently valued these items high and students also reported a positive mean score on this scale. On the naïve side, EFA separated items connected to the objective nature-of-knowledge from items connected to the objective nature-of-knowing. In line with our expectations, experts clearly rejected these items. In contrast, students on average tended to report a neutral position or even a limited agreement. Finally, EFA extracted two factors related to subjectivism. Contrary to our expectations, these factors did not follow our distinction between naïve and nuanced ideas. Furthermore, scale-reliability of these factors was low and expert answers showed a large variance. Consequently, these two factors were excluded from the questionnaire. The final questionnaire consisted of three factors focusing on (1) methodological criteria for generating historical knowledge, (2) the objective nature of historical knowing and (3) the objective nature of historical knowledge.

We compared differences between school tracks (HG and PU) to test the sensitivity of the questionnaire. As expected PU-students were on average a bit

more critical towards objectivist items, and valued methodological criteria a bit more, although effect sizes for all differences were small. Finally, we found a moderate correlation between the value of methodological criteria and individual interest and a weak correlation between criteria and history grade. With both objectivist factors, no correlations were found.

This study showed that epistemological beliefs could not simply be divided between a naïve and a nuanced cluster. The study also showed that notions about the subjective or interpretative nature of historical knowledge remained difficult to assess and interpret. In contrast, both objective factors yielded meaningful information and showed a conceptually interesting difference between experts and students. In the conclusion, we proposed to reconceptualize the scale that measures methodological criteria for generating this knowledge, as a separate (third) dimension addressing *historical method—critical thinking* (cf. Kuhn et al., 2000). We suggested that a rejection of items related to both scales measuring objectivism, combined with an appreciation of methodological criteria (critical thinking) might be indicative of nuanced epistemological beliefs and of an ability to "coordinate the objective and subjective dimension of [historical] knowledge" (Kuhn et al., 2000, p. 310). We concluded from our results that Dutch exam-students might be characterized as *procedural objectivists*—a position in which a knower believes that applying the correct procedures can produce objective knowledge about the past.

## GENERAL CONCLUSIONS

Hereunder, we will first discuss in which ways the explicit instruction of causal strategies and second-order concepts, as well as the explicit reflection on epistemological ideas influenced students' learning and the ability of students to *engage* in causal historical analysis. We will also discuss two additional aspects that affected the outcomes on the performance task. Finally, we will look back on

both theoretical models and discuss the insights they yield, when defining instructional approaches for teaching historical reasoning.

## EXPLICIT INSTRUCTION OF CAUSAL STRATEGIES AND SECOND-ORDER CONCEPTS

Defining and teaching specific causal second-order concepts and specialized vocabulary was an important addition of our studies. The various descriptions of historical thinking and reasoning primarily used the term "meta-concepts" or "second-order concepts" to refer to high-order historical thinking concepts (e.g. "causation", "progress", "evidence", "significance"). In these meta-concepts, epistemological aspects and strategy knowledge were often included, and the models did not explicitly mention concrete (lower-level) second-order vocabulary in their conceptualizations (cf. Lee & Ashby, 2000; Lévesque, 2008; Seixas & Peck, 2004; van Drie & van Boxtel, 2008; VanSledright & Limón, 2006). This is understandable because the models provided a generic description of historical thinking and reasoning. Elaborating a specific subset, such as causal historical reasoning, focused attention on the concrete conceptual and linguistic demands this reasoning posed on students.

Second-order language has also remained a hidden aspect in most empirical research on history education. Previous studies have mainly focused on the teaching and application of (sourcing) strategies, without giving much attention to the vocabulary that is connected to these strategies. Studies with a stronger focus on disciplinary literacy mainly focused on describing or fostering students' ability to structure historical accounts (e.g. Coffin, 2004, 2006; De La Paz & Wissinger, 2015; van Drie et al., 2015). The importance of developing students' concrete causal vocabulary has been primarily discussed in practitioner-oriented literature. In these articles, compelling pedagogical approaches were developed that have informed the teaching and learning activities in our studies (e.g. Chapman, 2003; Woodcock; 2005).

In this dissertation, we explicitly included the knowledge of causal second-order concepts and specialized vocabulary in our definition of causal historical reasoning (besides causal strategies and epistemological ideas). In the intervention studies, we explored how students acquired and applied both causal concepts, as well as strategies. The studies showed that the knowledge of causal concepts, as well as the knowledge of causal strategies increased in the explicit teaching condition, but not in the control group. We also found that the heuristics of students in the experimental condition were more domain-specific, and in their learner reports, these students more often reported to have learned about specific concepts or strategies.

In their essay-revisions, students in the explicit teaching condition integrated many elements that were indicative of applying second-order concepts and domain-specific strategies. First, these students included 'specialized' causal vocabulary. Furthermore, we found that students integrated both concrete historical events, as well as abstract historical concepts in their essays. These revisions indicated that students had understood the interrelatedness in historical explanations of concrete events, with more structural developments and phenomena (historical context). This finding differed from earlier studies in which students were shown to have difficulty in reasoning with more structural causes and tended to overemphasize the role of concrete events and persons (e.g. Carretero, Jacott, Limón, López-Manjón, & Leon, 1994; Halldén, 1997). Our studies also showed that looking for multiple causes was something that 11th grade students had already grasped. Already at pre-test, all essays contained multiple causes and the learner-reports of students from both conditions mentioned, "looking for multiple causes" as an important strategy.

We found it very useful to distinguish between the concepts and strategies (and epistemological ideas) involved in the meta-concept of 'historical causation', because it allowed us to define causal historical reasoning and establish domain-specific learning goals with greater clarity. Because strategies focus on

the question *how* to engage in causal historical reasoning and second-order concepts focus on developing the vocabulary to *express* this reasoning, the pedagogical approach towards developing both types of knowledge differs. However, in practice we also found both types of knowledge to be intrinsically linked. For instance, in the learner reports and answers to the heuristics prompt, students from the explicit group mentioned both the concepts, as well as *applying* these concepts to categorize causes. We found that students could only engage in strategies, when they possessed the conceptual apparatus and vocabulary.

We concluded, that our studies showed that explicit instruction and practice of causal concepts and strategies led to increased performance on the writing task. However, we also found that the effects were limited and we did not find a main difference in the quality of the essays between the experimental and the control group. The fact that our intervention lasted only three lessons might partly explain this limited effect, but our qualitative study also suggested other aspects that might have influenced the quality of students' written explanations. First, we will turn our attention to an aspect that had been explicitly addressed in the lesson-unit, namely students' epistemological beliefs. Subsequently, we will discuss two additional aspects that influenced the application of causal strategies and concepts in written text—namely, knowledge of causal historical text-structures and knowledge of working with historical sources and evidence. We content, that in history education these aspects should also be explicitly addressed.

## EXPLICIT ATTENTION TO EPISTEMOLOGICAL BELIEFS

Epistemological reflection on the nature of historical explanations was another important element in our explicit teaching environment. In the second study, we found that the value students attributed to criterialist beliefs, as well as the value attributed to subjectivist beliefs increased in explicit condition. This finding

implied that our intervention did indeed affect students on an 'epistemological level', although effect sizes were small. The increase in the value that explicit students attributed to methodological criteria was in line with our expectations. The increased value on subjectivism contrasted with our expectations and with the theoretical underpinnings of the BHQ—this questionnaire associated a higher value of subjectivist items with more naïve beliefs (Maggioni et al., 2009). However, this outcome might be explained from the fact that students had just worked for three lessons on tasks designed to visualize historical explanations as interpretative constructions. Nevertheless, the finding problematized the interpretation of results found on subjectivity. In our questionnaire study, we found that experts too, varied widely in their appreciation of subjectivism and that valuing subjectivism could reflect both a naïve stance, as well as a nuanced stance in which historical knowledge is seen as constructed and temporary. We concluded from our experiences that results related to subjectivism could not be adequately interpreted. This finding was in line with other studies employing the BHQ (e.g. Mierwald et al., 2016).

Although three lessons might have been too little to thoroughly influence students' epistemological beliefs, other results indicated that the epistemological dimension constituted a tangible and important aspect of the lesson-unit. Almost a quarter of the students in the explicit condition referred, in their learner reports, to having learned about history as an interpretation, for instance—"not everybody will consider the same causes as causes". Furthermore, the increased correlation in the explicit condition between criterialist epistemological beliefs and students' individual interest in history also suggested the importance of addressing the epistemological dimension in the classroom (cf. Alexander, 2005). In our questionnaire-study too, a correlation was found between individual interest and criterialist epistemological beliefs.

In the data, no significant correlations were found between task performance and students' beliefs about knowledge. Also in our qualitative

analysis, we could not define unambiguous relationships between epistemological beliefs and students' essays. In this study, we found that students from the experimental group on average integrated three causal claims in their revised essays. However, in spite of the relatively high mean value students placed on methodological criteria for constructing historical accounts, all essays lacked evidentiary and argumentative support for the claims students made. Furthermore, the 11th and 12th grade students in our final study on average reported a neutral position or even limited agreement towards objectivist ideas—a position that suggested that many students still expected historical knowledge to be fixed and dependent on 'correct' sources. This position might be related to our finding that most essays, even when they included analytic second-order language and causal claims, read as rather factual 'objective' reports. More fundamental research is needed to further explore the relationships between epistemological beliefs and historical reasoning.

In the end, we concluded that developing epistemologically rich learning environments and addressing epistemological questions related to causal historical explanations provided opportunities for making history relevant and interesting and constituted a 'tangible' aspect of the learning environment. However, we also found that influencing students' beliefs about historical knowledge and knowing remained a complex task and many methodological issues precluded the exact assessment of these beliefs. Finally, the relationships between epistemological beliefs about history and students' task performance remained unclear. It is possible that epistemological beliefs require additional knowledge—knowledge of causal strategies, but also knowledge of strategies for providing (counter-)arguments and working with historical sources and evidence—in order for students to be able to construct accounts that demonstrate their beliefs. These additional demands will be discussed in the next paragraph.

ADDITIONAL DEMANDS OF WRITING A HISTORICAL EXPLANATION

Both intervention studies problematized the ability of students to apply their knowledge of causal concepts and strategies in a writing task. Only on the criterion "integration and use of second-order concepts" did we find a significant difference between both conditions. Furthermore, correlational analysis yielded no stable relationships between the quality of students' written explanations and students' knowledge of causal reasoning strategies, second-order concepts and epistemological beliefs. These findings suggested that other aspects also influenced that quality of students written historical explanations.

Qualitative analysis revealed two additional demands that influenced outcomes on the performance task. The first demand was related to an integral aspect of our definition of causal historical reasoning. Although our lessons had primarily focused on students' knowledge of second-order concepts and strategies, our assessment of the quality of students' written explanations also focused on argumentation and use of historical sources and evidence. In our results, we found that even the more thematic essays lacked argumentation and support of claims with evidence. In the paragraph above, we connected this lack of argumentation to students' epistemological beliefs. However, another explanation might be that these strategies related to critical reasoning about historical sources should also be explicitly taught. This conclusion adheres to research that specifically focused on teaching these heuristics (e.g. Nokes et al, 2007; De La Paz, 2005).

Second, we found text-structure to be related to the amount and type of revisions that students included in their essays. The studies of Coffin have provided many insights in the domain-specific characteristics of different historical genres (2004; 2006). In this dissertation, we distinguished between a chronological linear structure and a thematic, nonlinear structure. Essays written within a more thematic structure were found to be more adaptive to integrating

analytical concepts and causal claims, whereas essays written within a more chronological structure focused more on 'organizing the facts'. However, although text-structure influenced task performance, we did not explicitly teach this 'genre-knowledge', and – as was witnessed by the hybrid character of most texts – students in general, possessed only implicit knowledge of these structures. In Dutch history-classrooms, writing is often regarded as a generic skill and historical writing is not included in the Dutch history-curriculum.

Based on our studies, we advocate that this knowledge of causal historical 'genres' and rhetorical demands, as well as working with historical sources, evidence and argument should receive explicit attention in lessons focussing on causal historical explanation. Apart from causal strategies and concepts, these aspects also influence the quality of students written explanations.

## EXPLICIT TEACHING, A CRUCIAL PRINCIPLE IN FOSTERING DOMAIN-SPECIFIC REASONING

The MDL allowed us to connect our definition to (generic) pedagogical principles. Three pedagogical principles were established that we expected would provide a necessary context for developing domain learning—(a) raising situational interest, (b) organizing social interaction, and (c) working on open and realistic tasks and questions. The results in the control group lent support to the conclusion that lessons based on these principles were able to elicit students' situational interest, support the acquisition of first-order knowledge, and led to enhanced performance in the post-test essays (perhaps caused by the increase of first-order knowledge). The learner reports also supported the idea that students in the control group had appreciated the inquiry-approach. However, judging by their learner reports and heuristics, these students appeared to have interpreted the inquiry task primarily as a generic form of inquiry and their answers did not portray an understanding of the discipline-specific nature of *causal historical*

inquiry. The main conclusion we drew from the studies in this dissertation was that explicit teaching of causal strategies and concepts indeed constitutes a crucial principle in order to achieve "rooted relevance" and "principled understanding", really develop students' ability to explain historical events. The studies provided rich evidence for the claim that this explicit teaching should focus on multiple aspects of (causal) historical reasoning.

## DEFINING CAUSAL HISTORICAL REASONING IN THE CONTEXT OF THE MDL AND THE FRAMEWORK OF HISTORICAL REASONING.

In this dissertation, we combined the framework of historical reasoning with the model of domain learning in order to define a domain-specific instructional framework for teaching and learning causal historical reasoning. Hereunder, we will briefly discuss how the MDL supported and appended the framework of historical reasoning.

The framework of historical reasoning was originally designed as an analytic framework to describe (and code) the cognitive, domain-specific components of historical reasoning—asking historical questions, constructing answers by using historical (first-order) concepts, meta-historical (second-order) concepts, contextualizing historical events and developments, providing (counter)arguments for claims and supporting these claims with evidence derived from historical sources (van Drie & van Boxtel, 2008). Because our research questions focused on designing effective instruction to foster causal historical reasoning, attention shifted from describing the constituent components of historical reasoning to defining the 'resources' that students required to engage in causal historical reasoning. The MDL differentiated between several cognitive and affective "resources", and therewith allowed us to define the types of knowledge involved in causal historical reasoning.

First, the MDL emphasized the importance of domain-knowledge in the development of expertise, and this aligned with the framework of historical reasoning that mentioned historical and meta-historical (or second-order) concepts as important components of historical reasoning. In line with this, we defined acquiring first- and second-order conceptual knowledge as important building blocks (or 'resources') for constructing causal explanations. Second, the goal to foster the ability to reason causally in history raised the question *how* (causal) historical analysis should be conducted. In the framework of historical reasoning strategy knowledge is not separately distinguished, but in the MDL developing (domain-specific) deep-level strategies is a core element of developing expertise. Therewith, the MDL added a focus on the knowledge of causal strategies to our definition. Third, the MDL added to the framework of historical reasoning through its focus on epistemological beliefs. This aligned with an increasing amount of research that investigates students' beliefs about historical knowledge and knowing (e.g. Buehl & Alexander, 2001; Maggioni et al., 2009).

Integrating the role of interest in our conceptualization of the ability to reason historically was a fourth addition of the MDL. In the MDL, raising (situational) interest was conceptualized as a precondition for engaging in deep-level strategies and developing a "rooted" understanding of the key concepts in a domain. In turn, an increase of domain-knowledge and deep-level strategies was also expected to induce higher levels of individual (or intrinsic) interest. Our second study confirmed that engaging students in causal historical reasoning and explicating concepts, strategies and epistemological underpinnings indeed stimulated their situational interest. In an updated version of the framework of historical reasoning, ideas about the nature and construction of historical knowledge, as well as students' interest have been included as "resources" that shape historical reasoning processes and outcomes (van Boxtel & van Drie, in press).

In addition, the framework of historical reasoning appended and deepened insights taken from the MDL. The framework provided an important starting point for defining the discipline-specific aspects of causal historical reasoning. The framework also suggested the interrelatedness of different components when answering a causal historical question. In line with the framework of historical reasoning our analysis of the essay-task indeed showed how the ability to apply second-order concepts and strategies in order to answer a causal question was influenced by students' ability to provide arguments and counterarguments, and work critically with historical sources.

Our studies showed how causal historical text-structure affected the type and quality of students' revisions. Based on this, we content that both generic and more domain-specific models of domain learning, should also pay attention to the ability to express higher-order domain-specific thinking and reasoning in writing. The framework of disciplinary literacy proposed by Goldman et al. (2016) is an interesting example of such an integrative model. In this framework, *discourse and language structures* are identified as one of the core constructs of disciplinary knowledge (besides, epistemology, strategies of reasoning, and overarching concepts). Based on our studies, we advocate that this discourse and language structures are operationalized both on the level of discipline-specific text-structures (or genres), as well as on the level of concrete second-order language and causal vocabulary.

## METHODOLOGICAL CONSIDERATIONS AND DIRECTIONS FOR FUTURE RESEARCH

Of course, the studies in this dissertation hold limitations. In this paragraph, we present some of these limitations and discuss consequences for future research. First, we will discuss the pedagogical considerations and subsequently reflect on the limitation of our measurement instruments.

PEDAGOGICAL CONSIDERATIONS

The first limitation of the two intervention studies is that we designed a learning environment that focused on multiple aspects of causal historical reasoning and integrated multiple pedagogical principles, learning activities and tasks. The pedagogical choices and domain-specific learning goals were conceptually linked through the use of the MDL, and the principles were elaborated with learning activities and tasks that previous research (primarily on history education) had shown to be effective. An important addition of our studies was that we (a) combined these activities and tasks in one coherent learning environment and (b) empirically investigated this learning environment on its ability to foster (causal) historical reasoning. A drawback of this approach is that we cannot pinpoint the effectiveness of separate elements of our design (e.g. developing graphical representations, modeling, group work, raising situational interest, epistemological reflection). We maintain that designing and investigating theoretically grounded learning environments constitutes an important direction of research, because it translates outcomes of educational research to realistic classroom contexts and connects more isolated findings within a larger framework of domain-specific instruction. Furthermore, this research allows asking (and answering) different questions, for instance about how pedagogical principles, or historical knowledge and beliefs, interact. Of course, future research should also continue to focus on isolated activities, tasks, and instructional approaches (such as working on graphical representations, modelling, or classroom discussion) in order to further our knowledge of the conditions under which these elements are most effectively applied.

A second limitation is that our intervention studies focused on constructing a causal explanation about one specific historical topic—World War I—, and that they were conducted on one school, and targeted one school-level and age group—11th grade preuniversity students. Limiting the sample in these

ways, allowed us to control the experiment tightly (e.g. training two experimenters to conduct the lessons and randomly assigning students to a condition) and therewith heightened the internal validity of the experiment and the ability to draw causal inferences. However, this choice also had consequences on the generalizability of the outcomes. The conclusion that our approach fostered students' causal reasoning ability is limited to this specific group of students and to the topic of World War I.

However, we expect our pedagogical principles in general, and explicit teaching in particular, as well as our conceptualization of causal historical reasoning to provide a valid framework, also when operationalizing learning environments that focus on different topics, age groups and school-levels. Of course, the choice for specific learning goals related to causal historical reasoning depends on the topic, age group and level—in a 7th grade class learning about the Peloponnesian wars in ancient Greece, we might limit our goals to understanding the concepts of cause and consequence, or to explicate multi-causality in history, whereas an 11th grade class studying the Weimar Republic, might focus on the concept of historical agency, on determining the significance of multiple causes, or on writing a causal argument. However, in each context, defining the knowledge of causal strategies, second-order concepts and epistemological understandings is important, sometimes in combination with aspects of historical writing, and/or working with historical evidence. Furthermore, we expect that open tasks, situational interest, making thinking visible, group work and explicit instruction are important in all lessons focusing on historical reasoning, although the lessons will vary in the amount of complexity and structure embedded in the learning tasks, as well as the amount of scaffolding the teacher provides.

Future research should focus on broadening our pedagogical approach—developing and investigating open tasks, lessons, and rubrics tailored for different age groups, educational levels, historical topics, and most importantly for other types of historical reasoning. In addition to theoretically

grounded studies, we suggest that it is also important to collaborate intensively with history teachers. In our intervention studies, we 'eliminated' the role of the history teacher and relied on two experimenters to teach the lessons. This decision was made in order to strengthen treatment fidelity, because our questions primarily focused on student learning. However, the next step would be to explore and strengthen history teachers' abilities to apply our pedagogical principles, as well as to support their explicit knowledge of (causal) historical reasoning. Previous research has found that although (aspiring) history teachers value historical inquiry and epistemologically rich learning tasks, they lack self-efficacy and find it difficult to translate their ideas to concrete lessons (cf. Voet & de Wever, 2016, 2017; Wansink, Akkerman, & Wubbels, 2016). An interesting question is how teachers can be supported to 'shift' their lessons towards a more explicit integration of (causal) historical reasoning in ways that are both practical and sustainable.

Our intervention studies aimed primarily at fostering the ability of students to construct a causal analysis. However, our essay task revealed that transferring knowledge of causal strategies and second-order concepts into a written explanation remained a complex challenge. In the writing tasks, other aspects played a role, such as the ability to work with historical sources, to underpin claims with historical evidence and arguments, and students' knowledge of causal historical text-structures. Future research should seek to develop pedagogical practices that supports students' ability to apply their knowledge of strategies and second-order concepts into (causal) historical writing. One suggestion for this research is to put more focus on deconstructing and analyzing causal explanations generated by peers or experts, in contrast to constructing own accounts. Being able to critically deconstruct historical accounts is an important historical competency (cf. Schreiber et al., 2006). In daily life students will often be confronted with the (sometimes-biased) accounts of others. A pedagogical argument for deconstruction is that writing research has shown that evaluating

(weak or strong) (peer-)models allows for a greater focus on strategies and therewith increases the development of explicit meta-knowledge (Hillocks, 1986; van Drie et al., 2015).

Our study found a stable correlation between interest and first-order knowledge. At the same time, we concluded that our lessons, with their focus on inquiry and epistemological questions stimulated students' interest. It would be interesting to design a longitudinal quasi-experimental study that investigates the effects of our approach on the development and retention of first-order knowledge, in relation to a more traditional textbook approach. We hypothesize that our approach will foster the acquisition of comparable amounts of first-order knowledge, but that retention of the key-concepts, developments and phenomena will be higher.

## LIMITATIONS OF MEASUREMENT INSTRUMENTS

Another limitation of our studies was that several instruments were developed specifically for this study (the causal-strategies-and-concepts questionnaire, the first-order knowledge test and the essay-task). An important strength of these instruments was their theoretical proximity to the different types of knowledge involved in causal historical reasoning. Nevertheless, this 'situatedness' might also hinder the applicability of the instruments with other topics, levels and age groups, or interpreting the outcomes across different contexts. This problem might be especially true for the writing task, because such a task will always differ depending on the topic. In our studies, we found that the scoring rubric provided a useful 'tool' to draw generic conclusions about the quality of students' (causal) historical writing. In the rubric, we defined the qualities of causal historical writing and the strategies and concepts we expected to find. Therewith, the rubric provided a domain-specific lens that might also be used to look at causal historical essays in general. Furthermore, the rubric might be used as an instructional tool

to give (peer-)feedback to students and inform curricular choices. Future research could strengthen the rubric, develop rubrics for other aspects of historical reasoning, and inform an educational curriculum that centers on a progression of (causal) historical reasoning. Such a curriculum could in turn lead to adapting and expanding the instruments developed in this study, and to calibrating the outcomes for multiple age groups and school levels.

Measuring students' beliefs about the nature of historical knowledge remained a difficult challenge throughout our studies. Although the alternative epistemological-beliefs questionnaire developed in de final study provided many insights, future research should seek to strengthen the stability of the questionnaire and enlarging the two scales related to objectivity. An important point of discussion in future research is the position of ideas related to subjectivism—should these notions be conceptualized as a separate scale, or does a rejection of objectivism indicate a move towards regarding historical knowledge as constructed? Conducting thinking-aloud studies with the goal of removing ambiguity and broadening the applicability of the questionnaire to a larger age group is an important step. Finally, questions remain about the relationship between epistemological beliefs and performance on historical reasoning tasks. Future studies should further explore these questions.

## IMPLICATIONS FOR PRACTICE

Several researchers have concluded that in history classrooms, historical reasoning remains mostly implicit (van Boxtel et al., 2010; VanSledright, 2011) and a study by van de Grift, van der Wal, and Torenbeek (2011) found that teachers in general find it difficult to foster higher-order skills. The first implication that we draw from the studies in this dissertation is that explicit attention to the multiple aspects involved in (causal) historical reasoning constitutes an indispensable principle if we want to foster the ability in students

to write historical accounts. Important background principles in this explicit teaching environment are open, historical questions and tasks, social interaction, making teacher and student thinking visible and raising situational interest.

In our studies, we defined design principles and elaborated a conceptual framework in which we operationalized (causal) historical reasoning. Based on this framework, we designed and investigated concrete learning activities, and developed rubrics and assessment instruments. These 'tools' allowed us to foster students' causal historical reasoning. It is our goal that the lessons and products developed in this dissertation can be used by history teachers and that the approaches we developed can support them to design their own explicit lessons aimed at fostering higher-order skills.

An example of this 'transfer to educational practice' is the professional development project "Beyond the facts", in which we work with a group of experienced history teachers (project number 405-16-508, Netherlands Organisation for Scientific Research). In the project, we analyze students' answers on causal questions and subsequently design tailored lessons based on our design principles (e.g. diagnosis of prior knowledge of causal reasoning, open tasks, explicit instruction, and social interaction). Preliminary findings indicate that this approach fosters teachers' ability to teach (causal) historical reasoning and adapt instruction to the needs of the students. These professional networks of history teachers, teacher educators and researchers on history education can constitute important 'spaces' in which we can develop a shared 'language' and explicate with greater clarity the strategic and conceptual knowledge that students in history should master. Within these networks, high quality lessons, rubrics, learning tasks and assessment instruments can be designed and adapted for specific levels and age groups.

Another implication for practice from our studies is that reflection on the nature and sources of historical knowledge deserves a more prominent place in the (Dutch) history classroom. In our studies, explicitly addressing

epistemological questions appeared to contribute to students' interest and the value they attribute to history. Our questionnaire study found that students on average hold mildly objectivist notions about history. Through centering history lessons more on historical questions and inquiry, students might begin to regard history not as a fixed representation of the past that authorities (e.g. teachers or textbooks) transmit, but as a debate in which they can participate, and in which evidence and arguments are the fundamental criteria. In line with Wansink et al. (2016) we advocate an increased attention to the pedagogical consequences of the interpretative nature of historical knowledge in teacher education.

Our studies emphasized the complexity of writing historical accounts based on the reading of multiple sources. Our intervention primarily focused on the strategies and concepts students needed to answer a causal historical question. However, when answering the document-based question, students were required to combine their (newly acquired) knowledge of causal strategies and concepts, and their (limited) first-order knowledge, with a critical analysis of historical sources, the rhetorical demands of providing arguments and counter-arguments and their (implicit) notions of historical text-structures. It is not surprising that students only partially succeeded in this task. We suggest that in history lessons these aspects are sometimes taught in isolation, but that students should also practice these skills in realistic (and thus integrated) tasks. These tasks should be relatively small and well designed, and detailed scoring rubrics are necessary in order to allow adequate scaffolding and well-focused feedback. A study of van Drie et al. (2015) showed that providing students with a one-hour domain-specific writing instruction enhanced their ability to translate historical reasoning to written text. In line with the MDL that advocates, "Teach more about less", we maintain that a focus on realistic and complex tasks should be accompanied with sharp choices regarding the amount of first-order content. We believe that, if we want to convince students that history is a subject in which you can progress (just like in mathematics), and not just a subject in which you reproduce

knowledge about consecutive periods and events, that historical reasoning and an inquiry-based curriculum are the key.