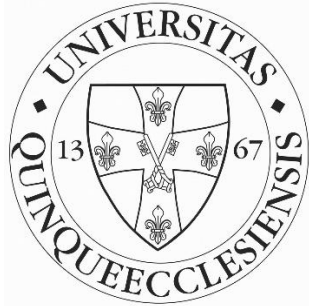University of Pécs

„Education and Society"

Doctoral School of Education

László Kojanitz

Analysis and assessment of intermediate baccalaureate history examination essays concerning the quality of their presentation of causality

Doctoral (PhD) dissertation

Thesis supervisors
Prof. dr. Ágnes F. Dárdai
Professor, PhD habil
Dr. József Kaposi
Associate professor, PhD habil

Pécs, 2017

Problem statement

Higher cognitive skills such as critical thinking, problem-solving skills and metacognition form the most valuable part of academic knowledge since the flexible adaptation of the acquired knowledge base cannot be realized without them (Wilson–Conyers 2016; Csapó 1998; Csapó 1999). Having such skills enable us to correctly interpret the phenomena in our environment as well as to find answers for the emerging problems (Endréné Réthy 1998). This development direction predominates the efforts aiming to renew the teaching of history. "The learning of history means owning the skills to become "memory-able", i.e., to acquire historical thinking. A key aspect of this thinking is that we reconstruct the past by using the competencies needed for understanding it, while having the appropriate methodological skills and the historically valuable sources from the past" (F. Dárdai, 2006: 5).

The conscious development of cognitive skills may also be the appropriate strategy for managing the problems arising from the constant growth of the knowledge base. Although we do not have the means to teach every key historical event, we can still provide our students with the intellectual tools that can help them understand historical events or problems at any point later on (Van Drie–Van Boxtel 2008; Van Sledright–Limon 2006). "Students need the best tools we can give them, understandings that enable them to think clearly about, for example, what kind of evidence is needed to support a particular kind of claim or what questions are being addressed in competing accounts." (Lee 2005: 70).

The development of historical thinking cannot be successful unless students' substantial and procedural historical knowledge is activated and improved concurrently. However, academic practice frequently demonstrates the typical problem that procedural knowledge is often developed only implicitly so this learning process remains almost completely hidden from students (Lévesque 2008). In connection with the above, the other problem is that assessment typically focuses on testing substantial knowledge only. So the development of historical thinking requires a renewal of testing culture as well, but it is quite difficult to grasp student performance fluctuations taking place in this area.

All testing processes must be based on three pillars: awareness of students' cognitive processes, observation of student performances and the interpretation of observations (Pellegrino, Chudowsky, & Glaser 2001). Thus the need for assessing historical thinking had researchers face three significant problems. What does historical thinking mean? How and with what exercises could we explore students' knowledge in this area? With regard to the

development level of historical thinking, how can we correctly interpret the results of the completed exercises?

The research projects conducted in the past decades have been instrumental in developing various cognitive models on the skills playing a role in historical thinking as well as the functioning and improvement thereof. Based on these cognitive models, we have developed and tested special assessment tools designed to determine the current level and improvement of the above skills (Ercikan, Seixas, 2015).

Research objective

With regard to the existing international results, I had various research areas to choose from. Of all factors and tools involved in the development of historical thinking, I eventually decided to investigate the assessment of students' school work. My choice was determined by two factors: firstly, it seems the hardest task to accomplish even at international level today and, secondly, because the targeted analysis of student skills and the renewal of assessment culture have a very significant impact on the teachers' work. So, if you want to conduct research that would affect school practice, it is advisable to especially focus on assessment tools. The teaching process cannot be renewed without having teachers confronted with the results of their earlier practices because that is the way to make them open for new methods.

In practice, the quality of historical thinking can primarily be assessed based on how a person can recall and successfully apply their relevant substantive and procedural knowledge needed to answer the questions emerging in relation with a new problem. Students' essays can especially reveal a lot about the maturity of their authors' historical approach and historical thinking. In order to assess that however, we would need to know exactly which content features of these texts may be considered as evidence and how they could be used for assessing the quality of historical thinking. However, we do not have such an assessment tool available to analyze student essays although it would be greatly useful for teachers in constantly monitoring and documenting their students' progress in historical thinking.

Without searching for and understanding the correlations of historical events, the mere knowledge of the past is worth next to nothing. The exploration and presentation of causal relations is what makes the difference between a chronicle-like recording of past events and historical narratives of pragmatistic historiography. The causal approach has been a key component of all, albeit greatly varying cognitive models (VanSledright, Frankes, 2000; Seixas, Peck, 2004; Van Drie, Van Boxtel, 2008; Lévesque 2008). For a long time, school history

teaching has considered it as a fundamental goal to develop students' skills necessary for creating explanations related to the reasons and consequences. That is why I decided to focus my research on causality of all components of historical thinking. My decision also derived from the key role of grasping causal relations in scientific and non-scientific interpretations and explanations of the past. It would be great to create an inner drive in students to constantly search for the reasons of events and changes taking place in the past and present.

So my research has been aiming to develop and test in practice a new system of assessment tools that could evaluate students' historical thinking including, in particular, their skills to represent causal relations, based on the historical analytical essays written by these students. The research included an overview of the findings of Hungarian and international research studies on the characteristics, development and assessment of historical thinking. The goal of this overview was to identify the content elements and features that are worth examining for such evaluation.

Preliminaries of the topic in professional literature and research

The characteristics of historical thinking

Historical thinking always means the comprehension, explanation and historical contextual positioning of an individual event and situation, but it also involves the application of general concepts and competencies allowing for the correct and professional investigation of various historical topics and problems. (Körber 2011; Taylor–Young 2003)

Historical science seeks answers for such questions and deals with such so-called wrongly-defined problems that have no one single solution. In fact, the solution of these problems can best be achieved by the amalgamation or competition of multiple interpretations with different aspects and methodologies. These interpretations help to better identify the limitations of how much a historical problem can be explored and reconstructed (Van Drie–Van Boxtel 2008; Voss–Wiley 2006).

The answers to the historical question represent primary conceptual knowledge. This answer can also be given by repeating what someone else has written down, even without independently interpreting or fully understanding the correlations lying in the explanations. A higher level of solving the problem may be achieved when the respondent has the adaptable conceptual knowledge for correctly interpreting and presenting the information related to the past (VanSledright & Limón 2006). The correct interpretation and application of historical

second-order concepts is a key component of this process (Lévesque 2008). However, the knowledge necessary for the correct interpretation and application of historical second-order concepts is closely connected to procedural knowledge as well. In the case of history, the above means a certain knowledge and applicability of methodologies and rules elaborated by historical science for the exploration and presentation of the past. These diverse activities may be arranged around the following three key tasks: acquire information; construct and analyze narratives as well as argumentation and problem solving (Voss & Wiley 2006).

So, as far as historical thinking is concerned, history teaching can only be considered to be having developmental value inasmuch as students have the opportunity to learn and try the methods and rules applied by historians since the background knowledge related to historical second-order concepts cannot be acquired without it (Kojanitz, 2011 and 2013). Equipped with such knowledge, students become more independent and their answers will reflect an increasingly problem-sensitive historical thinking. Without such cognitive interactions however, student responses will mean nothing more than the verbal recitation of lessons and we can never be sure if they truly understood what they seem to be speaking about so professionally. This is part of the reason for the demand to create such special assessment methods that could not only provide feedback on students' skills in terms of the content of responses but on the general quality of the approach to the problem and the explanations given in such responses, too.

The role and characteristics of causal explanations in historical thinking

In addition to presenting the causal relations, historical interpretations must also be able to show that several events and factors may have contributed in various ways to the occurrence of a particular historical event. The sophisticated presentation of this diversity requires the ability to identify various types of reasons. Interpretations based on causal relations include the presentation of conditions as well as reasons. In this regard, conditions mean everything that allowed the particular events or changes to occur while reasons mean the things that were necessary for these opportunities to occur. In this area, we can distinguish between manifested and latent events. The latter category includes such demographic and social or climate changes that can only be perceived on a long time scale (Spoehr–Spoehr 1994).

Jacott, López-Manjón and Carretero (1997) identify two theoretical models of explanation in history. The intentionalist model primarily focuses on human decisions, actions and activities. It attributes great significance to human motives, intentions and beliefs in terms

of how events unfold. In contrast, the explanations according to the structural model are based on the exploration of correlations between conditions determining social reality (for example: economic, demographic, social, political, religious conditions). In other models, the same idea is expressed by making a distinction between narrative and analytical causal explanations. The narrative type emphasizes human decisions and actions, presenting the causal relations in a chain-like sequence where one event causes the occurrence of another event. In contrast, analytical historical explanations focus on the mutual effects of underlying long-term social, economic and political processes. Consequently, this kind of approach to historical questions represents causal relations as complex networks much rather than sequences. (Coffin 2000)

The characteristics of the development of historical thinking

Jerome Brunner claimed that any subject can be taught at any age as long as we explicitly show students the internal structure of the given science, including its objectives, key concepts, analytical methods and the way it verifies and presents the findings of all the above. Brunner's suggested spiral curriculum thus requires us to return to the key concepts of the discipline again and again so that students could achieve an ever deeper understanding of them and their thinking could develop from the concrete to the abstract. (Bruner 1960)

Based on their studies conducted in the framework of the UK's Schools Council History Project (SCHP), researchers found that the appropriate task assignments were instrumental in arranging historical thinking skills into strategies and concepts that raised historical learning to a higher level. The latter were then defined as second-order historical thinking concepts. Historical change and historical reason were also identified as such historical second-order concepts (Shemilt 1980). According to the British researchers, these second-order concepts need to be interpreted correctly to enable students to understand, critically evaluate and interpret the new information and conclusions related to the past. "Learning history also requires an understanding of history as a discipline, evidenced in students' increasing understanding of key second-order concepts" (Lee, 2005: 69). The learning of second-order concepts can also be interpreted as the development of meta-cognitive historical thinking.

As the most important problem related to students' historical thinking, Lee identified the fact that, in their responses to historical questions, students interpreted such historical second-order concepts as history, past, cause, change, fact and truth in the same manner they use them in their daily lives. As he put it: "Students have ideas about the past, and about history, regardless   of

what and how we teach them." (Lee, 2005: 70) They create their own historical explanations with the help of the personal mini-theories according to which they imagine the world functions. Similarly to misconceptions in natural sciences, there is a real risk that the teacher fails to notice them, and even if the teacher calls the students' attention to this problem, the solution structures used by students in their daily lives will still continue to strongly influence their historical thinking. (Lee, 2005) In addition, Shemilt observed that students often interpreted causes as things that existed on their own and constituted some special category of events, and if there is a sufficient number of them, an event will occur, where the higher the significance of the event was, the more such causes it required to occur. So, students believe that if you want to find the answer to "why"-type questions, you need to identify and list as many of the above "cause"-type events as possible (Shemilt, 2000). A particular step towards the early onset of such mistaken thinking is when events existing on their own are represented as a linear causal chain in the explanations (Lee et al., 1997). However, this is still very far from the correlative system consisting of intentions, events, processes and circumstances as well as network-like mutual effects that could be the real solution for reconstructing past events.

When discussing the general tendencies characterizing students' historical thinking, researchers always point out the importance of individual differences between students (Lee, 2005). Researchers attribute these differences to various types of causes. In addition to content-related knowledge, general cognitive skills and age (Torney–Purta, 1994; Leinhardt–McCarthy Young, 1996; Perfetti et al., 1995; Wineburg, 1998), they also mention such factors as cultural background (Barton, 2001), epistemological beliefs (e.g., Kuhn et al., 1994) and memory capacity as well. Another important circumstance identified by the studies was that the level and development of understanding certain second-order concepts do not go hand in hand. In other words, a student who has a correct concept of historical changes may have significant problems understanding historical reasons. (Lee, 2005)

Students' relation to "why"-type questions goes through a permanent change due to their constantly growing practical experience of human behaviour and their learning of history (Kojanitz, 2015). Students need to acquire various types of explanatory principles and strategies to be able to correctly interpret the question related to the causal relations of past events and give adequate answers to them as well (Lee, 2001). In the first learning period, students' thinking is linear: they view each event as an inevitable consequence of the preceding events. As students' thinking develops, they realize that the progression of events is influenced by different kinds of mutually effective causes. As their historical consciousness improves, students also understand that causal relations are the unique combinations of various factors.

Eventually, they are able to understand causal relations as a co-effective network of multiple events and factors that are in a complex interaction with each other and they also realize that we can never fully learn the entire story. Instead of giving a chronological account of events, the presentation of such correlations rather requires students to create analytical explanations where the progression of events is presented based on the various pre-selected determining factors and correlations rather than just in the order in which they took place. In order to do so, they need to understand and interpret historical situations and problems in a much more abstract way, which means a significant jump in quality in terms of students' historical thinking (Coffin 2006).

The assessment of historical thinking

When designing the tools used for assessing the development of historical thinking, we must definitely consider the characteristics and progress of this special type of thinking, which does not make the challenge easier at all. First of all, there are typically more than one adequate answers for questions related to the past.

According to Chris Husbands (1996), tasks used for assessing students' thinking especially require teachers to very accurately pre-determine the objective and requirements of the tasks as well as the kind of solutions expected from students. The form and content parameters of the tasks (e.g., the length, the number, type and difficulty level of the sources involved in the task as well as the width and depth of the historical knowledge involved and how much the task can differentiate between students) must comply with the above characteristics.

In addition to the specifically-designed tasks, teachers also have other options for assessing the progress of historical thinking. These evaluations can also be based on such elements as the observations of students' class work, analysis of their written output and the teacher's documentation of the student's earlier performance (Husbands 1996). An insightful comment in class or a written homework may often be a much more important and reliable evidence for the development of the student's thinking than a final test designed for a summative assessment (Freeman, Philpott, 2009). However, it is not easy for teachers to adequately analyze the various types of written or verbal output of students from the aspect of historical thinking.

Criteria and sets of criteria used for assessing student responses and output can be differentiated from the aspect of whether they are more suitable for an analytical or a holistic

assessment of an output. The list created by Lomas (1990) is a good indication of the criteria that can be used for a holistic assessment.

Students can:

• move from the concrete and tangible towards the abstract and intangible;

• make a distinction between the different eras;

• summarize, categorize and generalize based on the details learned;

• create descriptions and explanations of past events and circumstances;

• identify evidence-based conclusions and evaluations and support them with argumentation;

• identify connections between the characteristic features and events of different eras;

• select historically significant topics and events, and present how they correlated with other important processes and changes;

• raise good questions and hypotheses and find the way to answer them;

• recognize the limitations of historical cognition,

• demonstrate their awareness and comprehension of inevitable uncertainties arising from the nature of historical knowledge.

The uneven development of historical thinking also poses a difficulty for assessors. The current performance of students may greatly depend on the type and even the topic of the task, too. Students' development also differs by the particular areas of procedural and conceptual skills required for historical thinking. If a good performance is registered in a particular area, it does not necessarily mean that the same applies to the other areas as well.

The difficulty of the challenge is shown by the fact that fully reliable and functional methodologies have not been developed for the assessment of historical thinking skills. Several empirical studies have demonstrated that experiments on the assessment of complex thinking have shown significant differences between the pre-determined cognitive goals and the exercises and assessments methods used for their evaluation (Baxtern Glaser, 1998; Ferrara, Chen, 2011).

Research questions and hypotheses

The research is based on the cognitive model below. The answer to a historical question represents primary conceptual knowledge. This answer can be produced by students simply recalling what they heard from the teacher or learnt from the textbook. This is possible even without students independently interpreting and correctly understanding the correlations indicated in the explanation.

A higher level of answering the question may be achieved when the respondent has the equally conceptual knowledge for correctly interpreting and presenting the information related to the past. The correct interpretation and application of historical second-order concepts is a key component of this conceptual knowledge. However, the knowledge necessary for the correct interpretation and application of historical second-order concepts is closely connected to the acquisition of procedural knowledge as well. In the case of history, the above means a certain knowledge and applicability of methodologies and rules elaborated by historical science for the exploration and presentation of the past. This diverse system of activities may be arranged around the following three key tasks: acquire information; construct and analyze narratives as well as argumentation and problem solving. So the idea of history classes is not to offer ready-made answers but to provide students with individually achievable tasks where conceptual knowledge and the development of procedural strategic knowledge are connected in various cognitive interactions. Equipped with such knowledge, students become more independent and their answers will reflect an increasingly mature and problem-sensitive historical thinking (Fig. 1).
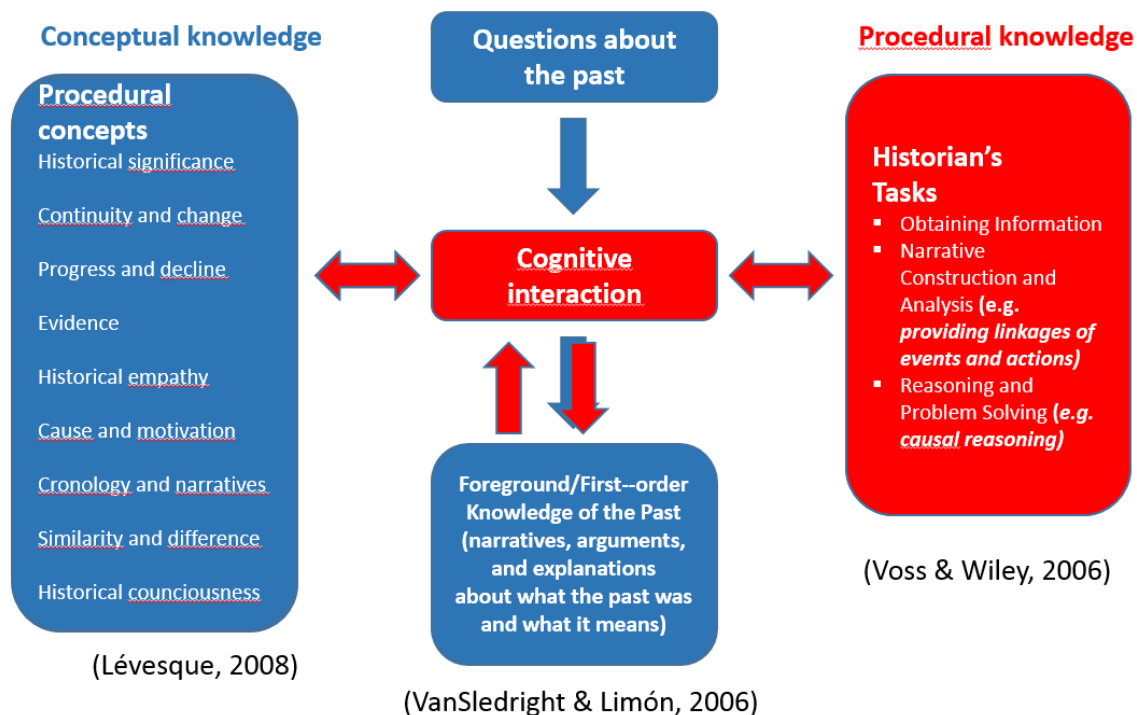
Conceptual knowledge

Procedural concepts
Historical significance

Continuity and change

Progress and decline

Evidence

Historical empathy

Cause and motivation

Cronology and narratives

Similarity and difference

Historical counciousness

(Lévesque, 2008)

Questions about the past

Cognitive interaction

Foreground/First--order Knowledge of the Past (narratives, arguments, and explanations about what the past was and what it means)

(VanSledright & Limón, 2006)

Procedural knowledge

Historian's Tasks
- Obtaining Information
- Narrative Construction and Analysis (e.g. *providing linkages of events and actions*)
- Reasoning and Problem Solving (*e.g. causal reasoning*)

(Voss & Wiley, 2006)

Fig 1. The conditions of cognitive interaction and its role in giving answers to historical questions

Higher knowledge can only be achieved through higher quality learning. The quality of learning depends on such factors as students' individual skills and interests as well as the methods and quality of history teaching. The practice of each teacher may significantly vary in terms of the quantity and quality of cognitive interactions instrumental in the development of historical thinking. The qualitative differences of students' knowledge, especially as far as the maturity of their historical thinking is concerned, are manifested in their papers as well.

The above also means that student essays written in response to "why"-type questions may greatly differ both in terms of elaboration and problem-sensitivity as well as by each individual or even by student groups. If that is so, then cognitive skills could be significantly developed through a special, important and useful assessment method which could provide feedback both for students and teachers on the content of students' answers as well as their approach to the problem and the quality of the explanations given in the answer. My research questions and hypotheses targeted such factors as the methodology of assessing historical thinking, the variances in student performances and the reasons thereof (Fig 5).

Research questions I.	Research hypotheses I.
Is it possible to assess historical thinking?	*Can we develop such assessment tools for analytical essays that could enable us to draw conclusions with regard to the quality of historical thinking in the individual students?*
What kind of cognitive model should we use for that?	The best foundations for the assessment of historical thinking may be the cognitive models that specifically feature the constant application and development of knowledge related to historical second-order concepts.
How applicable are they to real texts?	Analytical essays written in a real test situation may also be suitable for the assessment of historical thinking.
Can quantitative measurement tools be successfully applied for such assessment?	The joint application of quantitative and qualitative measurement tools improves assessment accuracy.
What criteria and what indicators do we need to apply for grasping the qualitative differences between the historical thinking of students?	Qualitative assessment can be successfully performed by applying assessment criteria and indicators that are based on the studies examining the development of students' historical thinking.
To what extent can we utilize the results of the historical thinking assessment for reforming and correcting the teaching-learning process?	The results generated by the assessment process can provide useful input for a differentiated development of historical thinking.
Research questions II.	**Research hypotheses II.**
How large differences can be identified based on the assessment of historical thinking?	*Are there significant differences between secondary school graduates?*
What differences are there between individual students?	Students have different skills and motivations, so they demonstrate significant individual differences in understanding and interpreting causal problems.
What differences are there between student groups?	The pedagogical practices of teachers vary in terms of the volume, intensity and quality of cognitive interactions, so student groups demonstrate significant differences of performance in understanding and interpreting causal problems.

Fig. 2 Research questions and hypotheses

Introduction of the research

The research focused on assessing the skills needed for the exploration and presentation of causal relations as well as for the causal explanation of historical events and changes. For the purposes of my research, I selected an analytical essay question that was part of the written secondary school graduation examination held in Hungary on 9th May, 2012. I decided that assessment tools should be first tested on analytical essays written in a real exam situation, because it allows for drawing realistic conclusions both on the practical applicability of these tools and the performance of the students as well.

Based on the four-dimensional characterization suggested by Martin Booth (Booth, 1992b), I made the following determinations about the task: The goal of the task: *Provide a causal explanation*; the difficulty level of the task: *Complex content and task*; the form and genre of the task: *Written explanatory essay*; the content of the task in terms of what prior knowledge and skills it requires from the student: *Almost completely builds upon the student's prior knowledge*

To design the assessment criteria, I used the cognitive model of Jannet van Drie and Carla van Boxtel (2008), which establishes that historical thinking is jointly defined by six cognitive skills: identify historical questions; use sources; place in context; argue, prove; use content related concepts; use second-order concepts.

After developing the assessment raster used for analyzing the students' exam essays, I primarily relied on such theoretical models and empirical studies that were related to the components and special developmental features of historical thinking. To identify the indicators, I used empirical studies examining the characteristics and development of students' causal historical thinking (Shemilt, 1980; Shemilt, 1983; Shemilt, 2000; Carretero et al., 1997; Jacott et al., 1998; Lee, 2005; van Drie & van Boxtel, 2008; Chapman, 2014). The above researchers used the clearly identifiable faults and deficiencies in the students' verbal and written output to characterize the different quality levels in their thinking. I used the same factors as indicators to identify and assess the qualitative differences of students' historical thinking as well (Fig. 3).

What typical deficiencies and faults occurred in students' causal thinking?
The simplified explanation of causal relations (e.g., putting events down to one single reason or human decisions and acts only, and/or considering events as a deterministic causal chain) Misconceptions: applying mini-theories drawn from their daily lives to explain historical events (e.g., treating unrealized options as irrelevant things or interpreting causes as some unique stand-alone events)

What characterizes a sophisticated causal thinking?
In terms of causal explanations, it prefers the structuralist model over the intentionalist model It analyzes and evaluates causes in their relative importance It places historical problems in a historical context It explains and reconstructs historical events as networks of correlations consisting of intentions, events, processes and circumstances as well as mutual effects.

Fig. 3 Determination of qualitative indicators used for assessing students' thinking

For the sake of completeness, I attempted to identify all such components and characteristics of essays on the causes of historical events that could be used as qualitative indicators for analyzing and assessing students' papers. In addition to the relevance of the content, I also considered as key assessment criteria such factors as the structure and linguistic formulation of the text, the author's sensitivity to problems and the interpretation of causal relations in terms of how abstract the presented causes and factors were.

The exam essays were primarily analyzed from the aspect of how the causal relations were presented. The texts were also subjected to quantitative micro-analyses and multi-aspect qualitative assessment to determine which of these analytical and assessment methods could be best used, even in teaching practice, for comparing and assessing students' historical thinking.

I identified six sub-criteria and thirty-six indicators in the assessment raster: content relevance (4 indicators); content structure (9); explanation of causal relations (5); problem sensitivity (7); errors in content (5); deficiencies in presenting the causes (8). I gave scores based on the presence or absence of indicators in the analysis of each essay.

Research findings

Methodological theses

Processing and adaptation of international research findings

In preparation for the research, I studied the Hungarian and international literature on the characteristics and progress of historical thinking. The resulting paper was published, providing useful assistance for researchers and teachers as well. During the study, I tested several new analytical criteria and assessment method that complied with the existing research findings on causal historical thinking. When developing the assessment methods, I also considered the findings of other studies with similar objectives.

The quantitative measurement and evaluation of content relevance

The quantitative analytical method I developed for assessing content relevance also helped me to measure, through objective data, a key qualitative difference in student essays. This key difference lies in how causal statements are supported by facts and explanations.

The analysis of content relevance in the various essays enabled me to detect and demonstrate much larger differences in quality than the ones reflected by the scores given by the teachers evaluating these exam essays. On the whole, the comparison of scores and the data of the relevant content elements suggested that evaluating teachers tended to level up the scores. Even if the number of relevant content elements was low, the teachers did not give low scores for task comprehension or for the exploration of event-shaping factors, either.

The quantitative measurement and evaluation of linguistic tools

I decided it was worth analyzing the use of linguistic tools in particular, because the enhancement of linguistic tools actively used for presenting causality closely correlates with the development of historical thinking. The more numerous and sophisticated expressions students use in differentiating between the effects of the factors shaping the events, the more they will develop the need and the skill to identify such differences in the exploration and explanation of the causes for certain events. This correlation was demonstrated by the findings of the study as well.

Qualitative assessment of essays through complex evaluation rasters

When I compared the findings of the quantitative measurement of relevant content elements with the assessment of the same elements through the evaluation raster, I found that the results were in line with each other. For example, only the essays that identified different types of causes (e.g., economic, political) had typically a higher-than-average number of relevant content elements.

Furthermore, analyzing the texts with various methods also enabled me to grasp the most typical problems that individual students had in terms of presenting causal relations. Therefore, it is worth finetuning the methods tested here because these types of assessment could enable us to more accurately reveal the strengths and weaknesses of individual students, thus allowing for a targeted and differentiated design of teacher instructions and practice tests for skill development.

Content related theses

Content relevance

The findings of the research suggest that student exam essays may demonstrate significant qualitative differences in terms of the skills necessary for giving causal explanations. Significant differences were detected between individual students as well as between the overall performance of the two different student groups involved in the research. One of the key qualitative differences between the essays was the number of facts and explanations supporting the relevant statements. I observed that the members of one student group laid much greater emphasis on presenting the explanations related to the causes, too. These essays contained a much higher number of such content elements, which also explains why the total number of relevant content elements was higher in this student group than in the other one. We can assume that this difference may primarily be attributed to the more conscious and efficient pedagogical work. The quantity and quality of teacher explanations on causality, and the modelling and conscious practice of producing such explanations with the students may significantly improve students' historical thinking skills as well as the quality of the responses they provide. It is worth conducting further research into this correlation.

The number of relevant content elements and the length of the essays did not demonstrate a correlation. Neither did I detect a difference of content relevance between the essays of girls and those of boys.

Linguistic tools

The comparison of individual students and student groups both showed significant differences in terms of using linguistic tools. The members of the student group with better content related performance were also much better in terms of the number and quality of linguistic tools applied. In addition to the more examples of discussing causal relations in a sophisticated and creative way, these students also used a larger vocabulary. This supports the assumption that sophisticated thinking about a particular content is in close correlation with a larger active vocabulary, and the conscious classroom development of the latter has a positive effect on students' thinking and performance as well.

Content structure

The number of different aspects based on which a person can categorize factors affecting historical events also has a significant role in how complexly this person can think about these factors and notice and present the various components of the problem. Of course, a well-, or not so well-structured special concept base plays a major role in the quality of the content of the given explanations.

The findings of content structure analyses demonstrated a significant difference between the two student groups in this regard as well. For example, the essays of the student group with a better overall performance often contained references to long- and short-term effects as well as explicit or implicit distinctions of foreign and domestic political or economic reasons. However, only the writers of the best essays determined in advance what aspects they were going to use for categorizing the factors affecting the historical event, in this case the Austro-Hungarian Compromise of 1867. This leads me to the conclusion that the presence of such structuring methods may be an important and relatively easily detectable qualitative indicator for assessing historical thinking.

Interpretation of causality

A key indicator of the maturity of historical thinking lies in whether students can perceive and present a particular historical situation in its real complexity or, on the contrary, they feel satisfied giving overly simplified answers to the questions related to causality.

Another useful approach to grasp the differences in the maturity of historical thinking was to assess student essays based on whether they were closer to the narrative or the analytical model of historical explanations. The narrative type emphasizes human decisions and actions,

presenting the causal relations in a chain-like sequence where one event causes the occurrence of another event. In contrast, analytical historical explanations focus on the mutual effects of underlying long-term social, economic and political processes. Consequently, this kind of approach to historical questions represents causal relations as complex networks much rather than sequences. (Coffin 2000)

The findings of the research showed that the vast majority of students applied the narrative model to present the causes and circumstances affecting the Compromise. Apart from a few exceptions however, none of these essays aimed to attribute and narrow the process down to a single string of events. In line with the nature of this particular historical event, students typically tried to present at least two plot lines of the interactive effects of the events and circumstances influencing the decision of the imperial court and the Hungarian leadership.

Few essays demonstrated a clear preference of the analytical approach over the narrative presentation of events. With regard to the maturity of their historical approach, these essays were the closest in quality to the sample answer written by the teacher, and we can assume that these authors had the appropriate analytical skills and tools to process causal historical problems on their own.

Sensitivity to problems

To assess how sensitive the essays were to historical problems, I selected criteria that were connected to the recognition and presentation of the controversial nature of the historical situation. The findings of the analysis demonstrated a significant difference between the two student groups in this regard as well. The group with the better performance had many students who could solve this task very well, presenting the Compromise as the result of concurrent events and factors that had opposing effective vectors at certain times.

The essays showed varying qualities in terms of depicting the controversial nature of the historical situation and the dynamics of the changes as well as students' ability to present the long-term factors that affected the events in addition to the direct causal relations involved.

Higher historical cognitive skills

From the list of indicators applied in the study, I specifically selected the ones that I considered especially interesting and important for assessing the maturity of historical thinking. I believe the occurrence of any of the indicators below in a student's written or verbal

explanation is evidence for the presence of a higher historical thinking in terms of causal type problems:

- Introduce with problem statement
- Present the controversial nature of the historical situation
- Present the circumstances exercising opposing effects
- Depict the dynamism, acceleration and deceleration or halt of the historical process
- Present the potential alternatives of the events that actually occurred
- Raise own questions
- Show that one particular reason would not have been sufficient to cause the events on its own
- Present causality based on the analytical model
- Present causes and consequences as a network-like system of factors that have parallel effects and mutually affect each other
- Verify the conclusion on the causes by hypothetically investigating what would have happened if the particular cause had not been there or had not taken place

These indicators enable us to relatively easily detect which students are already capable, based on the explanations they submitted, of a higher understanding and presentation of causal-type historical problems. Based on the above, we can then assume that their causal historical cognitive skills have reached a relatively high level. The presence or absence of these indicators was also a good basis for comparing the performance and cognitive skill levels of individual students and student groups (Fig. 4).
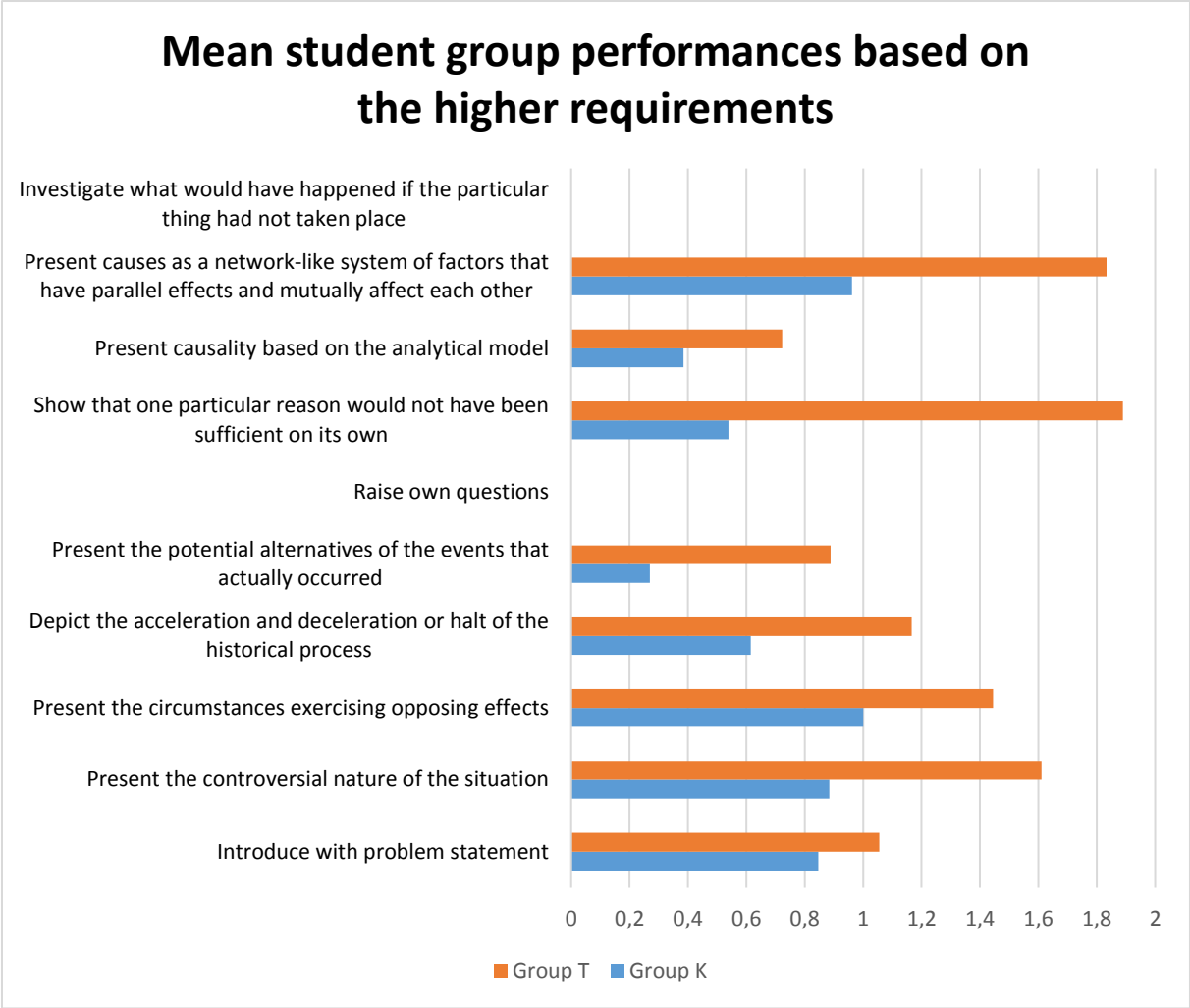
Mean student group performances based on the higher requirements

Category	Group T	Group K
Investigate what would have happened if the particular thing had not taken place		
Present causes as a network-like system of factors that have parallel effects and mutually affect each other	1.83	0.95
Present causality based on the analytical model	0.72	0.38
Show that one particular reason would not have been sufficient on its own	1.89	0.53
Raise own questions		
Present the potential alternatives of the events that actually occurred	0.88	0.26
Depict the acceleration and deceleration or halt of the historical process	1.16	0.61
Present the circumstances exercising opposing effects	1.44	1.00
Present the controversial nature of the situation	1.60	0.88
Introduce with problem statement	1.05	0.83

Fig. 4 The mean performance of the two student groups

Both student groups that were involved in the study consisted of highly skilled and motivated high school students. The significant differences detected between the results of the student groups support the assumption that the consciousness of the pedagogical work and the education methods applied by the particular teachers have a key role in the development of historical thinking. Of the above, the most important factors are perhaps how teachers model the explanation process of causes for historical events, whether they provide specific instructions and easy-to-understand qualitative criteria for writing such essays, whether students can practice these skills on their own frequently enough, and whether qualitative criteria similar to the ones in this study are applied with the appropriate weight, in addition to the assessment of factual knowledge.

Potential directions of further research

We are at such an early stage in the process of developing a special assessment methodology of the necessary skills for historical thinking, that we need preliminary studies like mine, i.e., studies to test the potential solution on a relatively small number of subjects. Meanwhile, we can already outline the plans and opportunities for the next steps. We should definitely provide the professional conditions for conducting regular studies targeting the historical thinking of students, the findings of which could then be considered for finetuning school practices. In order to do so, we will need to create theoretically sound and sophisticated development exercises and assessment tools.

Based on the findings of this study, one potential direction for further research could be to perform the assessment on a larger, representative sample of subjects. This could give us a picture of secondary school students' performance in terms of presenting causal relations today. Such a representative study could enable us to explore differences between individual students and student groups as well. The findings would allow for drawing useful conclusions regarding the efficiency of history teaching in secondary schools. In addition to students' knowledge of historical causes, we could assess their correct understanding and application of second-order concepts in history (e.g., historical change, historical source, historical evidence, historical interpretation). Such studies, even on their own, could call the attention of history teachers and professionals involved in developing content-related requirements to the importance of developing historical thinking. The findings of these studies would probably confront us with the typical deficiencies existing in the pedagogical work related to this area. Based on these findings, we could perhaps more easily convince the professional community that we need methodological reforms to achieve a more efficient development of historical thinking.

However, the studies on baccalaureate exam essays can provide only limited and indirect information on how deeply students actually understand the content they produce based on what they have learnt earlier. Furthermore, these essays do not allow for drawing exact conclusions as to what general level they have reached in terms of understanding the concept of and the problems related to historical causes. Based on the findings of the this study, we can only make an assumption as to which students could probably present independent opinions and conclusions and/or, relying on the correct criteria, make an independent judgement on the correctness and validity of the statements regarding historical causes. To study the above skills, we will need to analyze student essays written about the causes of events they have not studied before, which would require them to use the available information sources on their own or, as

a minimum, to be able to compare and evaluate different explanations. Therefore, we need to create such special exercises and sets of exercises that focus on the independent recognition, support and explanation of causality rather than the reproduction of explanations learned earlier. Furthermore, we also need to do productive research, with as many relevant results as possible, into the pedagogical work designed to develop historical thinking as well as the effects of the various exercises performed with the students.

The other potential research direction could be to apply the assessment tools tested in this study for examining the efficiency of new, experimental educational programmes and textbooks. We need to develop and test such new educational programmes, textbooks and digital education materials that could creatively adapt the international methods designed to improve historical thinking. In the event that these development projects are realized, we need to have reliable assessment tools to give us a realistic feedback on how these new teaching-learning methods and tools affect students' skills in comparison with control groups as well.

Another exciting potential research question is: what are the correlations between the knowledge elements and skills involved in the understanding and presentation of causal relations, and which of them have the largest influence on overall performance?

Assessment methods and tools must be constantly improved as well. In my case, this challenge may involve the following tasks:

- develop a similar assessment system as the one presented in my dissertation, in order to assess students' skills in relation with the understanding and explanation of other second-order concepts, e.g., historical changes, historical interpretation and historical significance;

- improve and finetune the developed assessment tools from such aspects as validity, reliability, objectivity and the interpretation of findings, and also improve the practical applicability of these assessment tools.

Further research must constantly strive for identifying connections to the ongoing research into the development of natural scientific thinking, as the cognitive models and interpretative frameworks applied in natural sciences may be useful for the research into historical thinking as well (Adey, Csapó, 2012).

References

Adey, Ph. és Csapó Benő (2012): A természettudományos gondolkodás fejlesztése és értékelése. In: *Tartalmi keretek a természettudomány diagnosztikus értékeléséhez*. Nemzeti Tankönyvkiadó, Budapest. 17−58**.**

Barton, K. C. (2001): A sociocultural perspective on children's understanding of historical change: Comparative findings from Northern Ireland and the United States. *American Educational Research Journal*, 38(4), 881–913.

Baxter, G., & Glaser, R. (1998). Investigating the cognitive complexity of science assessments. *Educational Measurement: Issues and Practices*, 17 , 37–45.

Booth, M. (1992a) *Students' Historical Thinking and the History National Curriculum in England*. Educational Resources Information Centre (ERIC), ED352 292.

Booth, M. (1992b): *How to Plan, Teach and Assess History in the National Curriculum*. Heineman Educational, Oxford.

Bruner, J. (1960), *The Process of Education*. Harvard University Press, Cambridge, MA.

Chapman, A. (2014): The 'Good Old Cause'? : Developing children's understandings of historical explanation. In M. A. Schmidt, I. Barca, A. C. Urban (Eds.), *Passados Possveis* (pp. 71-86). Ijui: Editora UNIJUI.

Coffin, C. (2000): *History as discourse: construals of time, cause and appraisal*. unpublished PhD thesis, School of English and Linguistics, University of New South Wales, Australia

Coffin, C. (2006) *Historical Discourse: The language of time, cause and evaluation*. London: Continuum.

Csapó B. (szerk., 1998): *Az iskolai tudás*. Osiris Kiadó, Budapest. 21-22.

Csapó B. (1999): Képességfejlesztés az iskolában – problémák és lehetőségek. *Új Pedagógiai Szemle*, 49. 12. sz. 4–13.

Dévényi A.– Gőzsy Z. (2014): Szempontok a történelemtankönyvek dekódolásához. In: Ünnepi tanulmányok F.Dárdai Ágnes tiszteletére, PTE, 2014. 145-166.

Ercikan, K., Seixas P. (2015):"Introduction: The New Shape of History Assessments In: *New directions in assessing historical thinking*, Publisher: Routledge, Editors: pp.1-15

F. Dárdai Á. (1997): Történelemdidaktikai koncepciók az újabb német szakirodalom tükrében = *Történelempedagógiai Füzetek* 2. 1997. 103-122.p

F. Dárdai Á. (2006): *Történelmi megismerés –Történelmi gondolkodás I-II*. Pécs, 2006. 14-29. p

F. Dárdai Á., Kaposi J. (2008): A problémaorientált történelemtanítás és a fejlesztőfeladatok In: Bánkuti Zsuzsa, Lukács Judit (szerk.) *Tanulmányok az érettségiről. Hatásvizsgálat, tantárgyi vizsgák értékelése, feladatfejlesztés.* 416 p. Budapest: Oktatáskutató és Fejlesztő Intézet (OFI), 2008. pp. 353-369.

Ferrara, S., & Chen, J. (2011*). Evidence for the accuracy of item response demand coding categories in think aloud verbal transcripts.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Freeman, J., Philpott, J. (2009): 'Assessing Pupil Progress': transforming teacher assessment in Key Stage 3 history. *Teching History*, 137. sz.

Husbands, C. (1996): *What is History Teaching? Language, Ideas and Meaning in Learning about the Past.* Buckingham, 1996.

Jacott, L., López-Manjón, A., & Carretero, M. (1998). Generating explanations in history. In J. F. Voss, & M. Carretero (Eds.) *Learning and reasoning in history. International review of history education (Vol. 2)* (pp. 294–306). London: Woburn.

Kaposi József (2015): Történelmi gondolkodás és a problémaorintált tanítás. In: Szujo Béla (szerk.) *Válogatott tanulmányok II.: Tanterv, Történelem, Módszertan* Budapest: Szaktudás Kiadó Zrt., 2015. pp. 9-27.

Knausz I. (2001): A történelemtanítás funkcióiról. In: Donáth Péter – Farkas Mária (szerk.): *Filózófia – Művelődés – Történet* 2001. Budapest, 2001, ELTE TÓFK.

Kojanitz L. (2011): A forrásfeldolgozástól a kutatásalapú tanulásig. *Történelemtanítás*, 4. sz.

Kojanitz L. (2013): A történelmi gondolkodás fejlesztése. *Iskolakultúra*, 2013/2.

Kojanitz L. (2015): A diákok gondolkodásának fejlődése a történelemtanulás eredményeként. *Iskolakultúra*, 2015/11.

Kuhn, D., Winestock, M., Flaton, R. (1994): Historical Reasoning as Theory-evidence Coordination. In: *Cognitive and Instructional Processes in History and the Social Sciences*. Eds. CARRETERO, Mario – VOSS, James F. Hillsdale, NJ., 1994. 377–402.

Lee P.J., Dickinson, A., Ashby, R. (1997): "Just Another Emperor": Understanding Action in the Past. *Internationla Journal of Educational Research* 27 (1997) 233-244.

Lee, P. J. (2001): History in an Information Culture. *International Journal of Historical Learning, Teaching and Researc*h, 1 (2). Retrieved August 21, 2009, from: http://centres.exeter.ac.uk/historyresource/journal2/journalstart.htm.

Lee, P.J. (2005) 'Putting principles into practice: Understanding history' in Donovan, M.S. and Bransford, J.D. (eds), *How Students Learn: History in the Classroom*, Washington DC: National Academies Press, pp.86. (85.)

Leinhardt, G., & McCarthy Young, K. (1996). Two texts, three readers: Distance and expertise in reading history. *Cognition and Instruction*, 14(4), 441–486.

Lévesque, S. (2008): Thinking Historically: Educating Students for the Twenty-first Century. University of Toronto Press, 2008.

Lomas, T. (1990): Teaching and Assessing Historical Understanding, Teaching of History Series, no 63, Historical Association, London.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment.* Washington, DC: National Academy Press

Perfetti, C. A. – Britt, M. A. – Georgi, M. C. (1995): *Text-based Learning and Reasoning. Studies in History*. Hillsdale, NJ., 1995.

Réthy Endréné (1998): Az oktatási folyamat. p221-270 In: Falus, I. (szerk.): *Didaktika: Elméleti alapok a tanítás tanulásához*. Nemzeti Tankönyvkiadó, Budapest

Kadriye Ercikan, Peter Seixas, (2015) In: *New directions in assessing historical thinking*, Publisher: Routledge, pp.17-29

Seixas, P. – Peck, C. (2004): Teaching Historical Thinking. In: *Challenges and Prospects for Canadian Social Studies*. Eds. SEARS, A. – WRIGHT, I. Vancouver, 2004. 109−117. (http://www.culturahistorica.es/seixas/ seixas_peck.pdf) [2013.03.02.]

Shemilt, D. (1980): *History 13–16 Evaluation Study*, Holmes McDougall, Edinburgh.

Shemilt, D. (1983). The devil's locomotive. *History and Theory*, 22(4), 1–18.

Shemilt, D. (2000). The Caliph's coin: The currency of narrative frameworks in history teaching. In P.N. Stearns, P. Seixas, and S. Wineberg (Eds.), *Knowing, teaching and learning history*. New York: University Press.

Spoehr, K. T. – Spoehr, L. W. (1994): Learning to Think Historically. *Educational Psychologist*, 29. (1994):2. 71–77.

Torney-Purta, J. (1994). Dimensions of adolescents' reasoning about political and historical issues: Ontological switches, developmental processes, and situated learning. In M. Carretero, & J. F. Voss (Eds.) *Cognitive and instructional processes in history and the social sciences* (pp. 103–122). Hillsdale, NJ: Erlbaum.

Vajda B. (2009)**:** *Bevezetés a történelemdidaktikába és a történelemmetodikába*. Selye János Egyetem, Komárom, 2009.

van Drie, J., van Boxtel, C. (2008): Towards a Framework for Analyzing Students' Reasoning about the Past. *Educational Psychology Re*view, v20 n2 p87-110 Jun 2008

VanSledright, B.A – Frankes, L. (2000): Concept- and Strategic-knowledge Development in Historical Study. A Comparative Exploration in Two Fourth-grade Classrooms. *Cognition and Instruction*, 18.: 2. 239–283.

VanSledright, B.A., Limon, M. (2006): "Learning and Teaching Social Studies," in Patricia Alexander and Philip Winne (Eds.), *The Handbook of Educational Psychology*, 2nd Edition (Mahweh, NJ: Lawrence Erlbaum Associates, 2006, pp. 545–570). Mahweh, NJ: Lawrence Erlbaum Associates.

Voss, J. F., & Wiley, J. (2006). Expertise in history. In N. C. K. A. Ericsson, P. Feltovich, & R. R. Hoffman (Ed.), *The Cambridge handbook of expertise and expert performance* (pp. 569–584). Cambridge, UK: Cambridge University Press.

Wilson, D., Conyers, M (2016): *Teaching Students to Drive Their Brains: Metacognitive Strategies, Activities, and Lesson Ideas*. Alexandria, ASCD

Wineburg, S. (1998). Reading Abraham Lincoln: An expert/expert study in the interpretation of historical texts. *Cognitive Science*, 22, 319–346.

Wineburg, S. (2001). *Historical Thinking and Other Unnatural Acts: Charting the Future of Teaching the Past*. Philadelphia: Temple University Press.

Publications related to the research topic

Kojanitz L. (1992a): A történelmi kultúra fejlesztése. In: Zsolnai József*: A magyar közoktatás minőségi megújításának szakmai programja. Iskolakultúra*, 1992/6-7.

Kojanitz L. (1992b): A tevékenység- és képességcentrikus történelemtanítás. *Iskolakultúra,* 1992/11-12)

Kojanitz L. (2010): A kérdésorientált (inquiry based) történelemtanítás összekapcsolása az IKT adta lehetőségekkel. *Iskolakultúra*, 20. 9. sz.

Kojanitz L. (2011): A forrásfeldolgozástól a kutatásalapú tanulásig. *Történelemtanítás*, 4. sz.

Kojanitz L. (2013): A történelmi gondolkodás fejlesztése. *Iskolakultúra*, 2013/2.

Kojanitz L. (2013): A történelmi kulcsfogalmak. *Taní-tani Online*, 2017. 03. 03-i megtekintés, 2017. május 12.) http://www.tani-tani.info/a_tortenelmi_kulcsfogalmak

Kojanitz L. (2015): A diákok gondolkodásának fejlődése a történelemtanulás eredményeként. *Iskolakultúra*, 2015/11.

Lectures related to the research topic

Kojanitz L. (2013): A kerettanterv változásai

Történelemtanárok Egyletének 23. országos konferenciája

Budapest, 2013. október 12.

Kojanitz L. (2017): A történelmi gondolkodás fejlesztése és értékelése

XVII. ORSZÁGOS NEVELÉSTUDOMÁNYI KONFERENCIA

Nyíregyházi Egyetem, 2017. november 9–11.