

### On the Way to an Internationally Shared Assessment of Historical Thinking (?)

**Veranstalter:** History Education Subject Group, Faculty of Education, Universität Hamburg Prof. Dr. Andreas Körber

**Datum, Ort:** 01.07.2016–03.07.2016, Hamburg

**Bericht von:** Patrizia Seidl / Andreas Zuckowski, Arbeitsbereich Didaktik der Geschichte, Universität Hamburg

Focusing on historical thinking and its assessment is an international trend in History Education.<sup>1</sup> At first sight, it is quite clear that different cultural backgrounds bring along different approaches of assessment: curricula, administrative and societal requirements, research traditions etc. differ substantially between countries. But when taking a closer look – especially at the different concepts and understandings of historical thinking – overlapping areas and connecting points become visible. This is especially the case when comparing research from Sweden (and Scandinavia), German-speaking countries, the Netherlands (and Belgium), Canada, the United States and the United Kingdom.

To secure this common theoretical foundation and to explore the options for future collaboration, researchers from seven different countries were being invited to Hamburg (Germany) to comparatively discuss current conditions, approaches and methods of history assessment. In his conference opening, ANDREAS KÖRBER (Hamburg, Germany) pointed out three arguments for collaboration in a cross-national development of assessment rationales and instruments: a) the post-traditional, heterogeneous character of many present societies, calling into question the „natural“ reliance of basing tests on a single, given „master“ narrative (and case knowledge based on it); b) the increased interest in comparative assessment of History Education outcomes across national borders; connected to this c) an increased demand of individual critical thinking, not only in national frameworks, but also on an international and global scale. According to Körber, the questions on the way to an internationally shared assessment of historical thinking can only be clarified partly by theoretical reflection. In

fact, it requires cross-examination of conditions, approaches, methods and criteria developed in different countries under different curricular conditions, with reference to different theoretical foundations.

Hence, the first section of the conference was meant to give an overview of the current understanding of history teaching, learning and assessment, concentrating on four specific regions. PER ELIASSON (Malmö, Sweden) elaborated on the development of schemes for the assessment of history learning outcomes in Scandinavia. For Sweden, he focused on the current (2011-) „standard“-related approach, pointing out that the Swedish „standards“ combine both: „competencies“ and content, with a congruent progression logic. After a „tour d’horizon“ of the present forms of assessment in the Nordic countries (Finland, Sweden, Norway and Denmark), Eliasson especially pointed out the need for more concrete formulations of standards (integrating content and skills) and de-central application of them (promising more impact on historical learning).

CARLA VAN BOXTEL (Amsterdam, Netherlands) presented the assessment of historical thinking and learning in the Netherlands and Flanders.<sup>2</sup> Starting with an overview of the Dutch debate about assessing historical knowledge versus assessing historical skills, she stated that the debate has shifted towards a more knowledge-related syllabus and assessment. While examples for the assessment of skills and the demand to apply genuine historical thinking-concepts can be found (e.g. *causation, continuity and change*), van Boxtel nevertheless concluded that in the Netherlands and Flanders most history assessment questions are designed to rather test historical content knowledge than historical competencies and historical thinking dimensions.

LINDSAY GIBSON (Alberta, Canada) started with a comparison of the two large

---

<sup>1</sup> The most prominent publication in this regard is certainly Kadriye Ercikan / Peter Seixas (Eds.), *New Directions in Assessing Historical Thinking*. New York 2015.

<sup>2</sup> Karel van Nieuwenhuysse from the Catholic University of Leuven (Belgium) kindly supported the Flanders-part of the presentation.

assessment programs enforced in the United States: the National Assessment of Educational Process (NAEP), being designed to track longitudinal trends and focusing on both, historical content knowledge and skills, and the Advanced Placement History Exams (AP), which are designed to provide high school students with college level coursework. Then, after briefly outlining the „Beyond the Bubble“ assessment project by the Stanford History Education Group, he focused on a specific Canadian approach. This approach consists of a one-hour test, explicitly designed to assess historical thinking and focusing on three of Seixas' & Morton's *Big Six* Historical Thinking Concepts.<sup>3</sup>

Closing the section, MONIKA WALDISWEBER (Aarau, Switzerland) and JOHANNES MEYER-HAMME (Paderborn, Germany) presented various examples and experiences with historical learning assessment in German-speaking countries. They exemplified competence-focusing tasks in Bodo von Borries' approach to measuring historical consciousness<sup>4</sup> and elaborated on Peter Gautschi's research on good history teaching.<sup>5</sup> Afterwards, Waldis-Weber and Meyer-Hamme introduced the cooperation project „Historical Thinking – Competencies in History (HiTCH)“.<sup>6</sup> In this project, a large-scale test for history assessment of 15-year old students could successfully be developed and evaluated from 2012-2015.

Altogether, the region-specific overviews show approaches for (large-scale) history assessment in all countries represented at the conference. Nevertheless, testing historical knowledge still plays a central role, especially when it comes to state-mandated testing. „Freelancing“ assessment projects show a clearer focus on the assessment of historical thinking abilities and skills.

The second section of the conference went more into detail, focusing on actual research and development projects on history learning assessment. Here, mostly open formats and open-ended tasks were the method of choice. CECILIA AXELSSON YNGVÉUS (Malmö, Sweden) uses qualitative interviews for her current research, dealing with school placement assessment. Within her project, she focuses on historical knowledge and abilities

of newly arriving students – most of them being refugees – to the Swedish school system. With regard to the presumably very different prior knowledge and experience, this program of assessment takes the form of „charting“ or „mapping“. This approach promises highly valuable insights and examples for curriculum and background-unspecific testing.

Also dealing with the assessment of knowledge and skills at the same time – but with a very different approach – WALTRAUD SCHREIBER, together with MICHAEL WERNER and MATTHIAS HIRSCH (Eichstätt-Ingolstadt, Germany), focused on how to relate logfiles from an award-winning multimedia textbook (the „mBook“) to competence test results.<sup>7</sup> Logfile data from mBook-learning lessons will be collected from about 2000 users (grade 9 and 12) who also regularly take part in a competence test (based on the HiTCH project) as well as a curriculum-based content knowledge test. Schreiber, Werner and Hirsch highlighted that the aim of the project is to identify correlations between the mBook usage and successful historical learning processes.

Another digital approach, rather focusing on methodological questions, was presented by CHRISTIANE BERTRAM and ZARAH WEISS (Tübingen, Germany). Their research

<sup>3</sup> Peter Seixas / Tom Morton, *The Big Six: Historical Thinking Concepts*, Toronto 2013.

<sup>4</sup> Magne Angvik / Bodo von Borries (Eds.), *Youth and History: A Comparative European Survey on Historical Consciousness and Political Attitudes among Adolescents*. Vol. A: Description. Hamburg: Körber Stiftung 1997; B. von Borries, *Das Geschichtsbewußtsein Jugendlicher: Erste repräsentative Untersuchung über Vergangenheitsdeutungen, Gegenwartswahrnehmungen und Zukunftserwartungen von Schülerinnen und Schülern in Ost- und Westdeutschland*. Jugendforschung. Weinheim 1995.

<sup>5</sup> Peter Gautschi et al. (Eds.), *Geschichtsunterricht heute: Eine empirische Analyse ausgewählter Aspekte*, Bern 2007.

<sup>6</sup> Ulrich. Trautwein et. al., *Kompetenzen historischen Denkens erfassen: Konzeption, Operationalisierung und erste Befunde des Projekts „Historical Thinking – Competencies in History“ (HiTCH)*. 2017 (in preparation).

<sup>7</sup> The „QQM project“ is a cooperation project of KU Eichstätt with the Institute for Digital Learning (Florian Sochatzy), Universität Lüneburg, Machine Learning Group (Ulf Brefeld) and Universität Tübingen, Empirical Educational Research (Ulrich Trautwein).

project (dealing with the events of October 9th 1989 in the former GDR) aims at computer-based rating and coding of open-ended tasks, based on computational linguistics data. First results of complexity analyses as well as automated content assessment were presented. Even though promising, main challenges lie in combining the two strands, and in checking the compliance of the results with narrative and empirical plausibility.

Still using open formats, but shifting from narrative skills and procedural competencies to the assessment of conceptual understanding, CARLA VAN BOXTEL (Amsterdam, Netherlands), gave an insight into her current research (together with Uddhava Rozen dal). The project aims at the development of tasks for formative assessment of historical causal reasoning. Van Boxtel introduced her updated cognitive conceptualization model of historical reasoning, differentiating three dimensions (epistemological beliefs, first- and second-order knowledge).

Adding to this, LISE KVANDÉ (Norwegian University of Science and Technology Trondheim, Norway), introduced an approach of fostering historical thinking and empowering students through conceptual learning. Drawing on the concepts and methods of the CLEAR research project (2010-2015), Kvande sketched an empirical approach to tracking the development of students in the usage of key concepts in mainly focus-group discussions. Her study uses a longitudinal design, surveying students in secondary school (2015-2018).

Most projects presented at the conference were aiming at the assessment of students in secondary school. MONIKA WALDIS WEBER and MARTIN NITSCHÉ (Aarau, Switzerland) added another target group by presenting their research project „VisuHist“ (2013-2016), which focuses on the assessment of professional knowledge and historical thinking of future history teachers. Based on the FUER model of historical thinking<sup>8</sup> (and integrating aspects from other cognition models)<sup>9</sup>, they analyzed historical narratives prepared by student teachers in a two-dimensional design. This research and follow-ups promise valuable insight into chances (and limits) of developing

cross-curricular assessment-tests.

Two presentations had a clear focus on large-scale assessment. PER GUNNEMYR and DAVID ROSEN LUND (Malmö, Sweden) presented a newly developed large-scale test for Swedish Upper Secondary Schools. The test is based on the Swedish curriculum and designed to support teachers in assessment and grading. A special feature of the open-answer format (in optional use since 2015) lies in the approach of granting the participating students influence on the selection of content when it comes to applying historical skills.

ANDREAS KÖRBER (Hamburg, Germany), contributed a comparative approach for the assessment of procedural historical competencies. After defining the term „historical thinking“ and presenting a spectrum of historical learning outcomes, he introduced, analyzed and compared different international approaches for historical thinking and historical competencies. Finally, he shared experiences of the HiTCH research group with developing principles and tasks for assessing procedural competences of students and discussed some task examples.

Summarizing, the presentations of the conference helped to identify a series of different aspects which need to be taken into account and to be explored by conceptual as well as empirical (piloting) studies when approaching international and intercultural assessment of history learning outcomes (e.g. the imperative application of open-ended and essay-formats beside closed formats). Furthermore, some projects were identified which might provide for specific approaches to address some of these questions, such as „bottom-up“-experience in identifying concepts and criteria for rating open-ended tasks, as well as specific methods for establishing cross-cultural validity and reliability in focused pi-

---

<sup>8</sup> Andreas Körber / Waltraud Schreiber / Alexander Schöner (Eds.), *Kompetenzen historischen Denkens: Ein Strukturmodell als Beitrag zur Kompetenzorientierung in der Geschichtsdidaktik*. Neuried 2007; Andreas Körber, *Historical consciousness, historical competencies – and beyond? Some conceptual development within German history didactics*, 2015.

<sup>9</sup> Jannet van Drie / Carla van Boxtel, *Historical Reasoning: Towards a Framework for Analyzing Students' reasoning about the Past*. *Educational Psychology Review* 20 (2008), pp. 87 – 110.

lot studies.

However, the conference also highlighted that in the face of concepts as well as assessments of History Education, several premises seem to be generally shared (such as focusing on students' own thinking and arguing). Still, a number of aspects will require further clarification and elaboration. Among them is e.g. the differentiation of levels of competences, the integration/distinction of several specific skills (such as „reasoning“ vs. „narrating“) and – maybe first to be addressed – the question of varying terminology in different languages, even in generally similar approaches (e.g. „skills“ vs. „competencies“).<sup>10</sup> Also, in the long run, a mere concentration on Western countries, traditions and research does not seem to be „well-balanced“ and will not be satisfactory. But if the conference participants were able to initiate a stable exchange and working process, this could be a starting point for a further internationalization of assessment in History Education.

## Conference Overview:

### Opening

Andreas Körber (Universität Hamburg, Germany)

### Section I: Region-specific (Pre)Conditions for Historical Learning Assessment

Per Eliasson (Malmö Hogsköla, Sweden): Scandinavia

Carla van Boxtel (University of Amsterdam, Netherlands): Netherlands and Flanders

Lindsay Gibson (University of Alberta, Canada): Canada and the US

Monika Waldis Weber (FHNW Aarau, Switzerland) & Johannes Meyer-Hamme (University of Paderborn, Germany): German-speaking countries

### Section II: Current Research and Development Projects on History Learning Assessment

#### Panel I: Assessment between Knowledge and Competencies (Mod.: Nicola Brauch)

Cecilia Axelsson-Yngvéus (Malmö Hogsköla, Sweden): Assessing the Knowledge and Abilities of Newly Arrived Students in the

#### Swedish School System

Waltraud Schreiber (with Matthias Hirsch and Michael Werner, KU Eichstätt, Germany): Relating mBook-Logfiles to Competence Test Results (QQM project)

#### Panel II: Assessing Procedural Competencies: Open Formats (Mod.: Lindsay Gibson)

Monika Waldis Weber / Martin Nitsche (FHNW Aarau, Switzerland): Assessing Narrative Competence with Open Formats („VisuHist“ project)

Christiane Bertram / Zarah Weiß (University of Tübingen, Germany): Computer-Based Evaluation of Student Texts

#### Panel III: Assessing Procedural Competencies: Large-Scale (Mod.: Béatrice Zielger)

David Rosenlund / Per Gunnemyr (Malmö Hogsköla, Sweden): A New Large-Scale Test for Swedish Upper-Secondary School

Andreas Körber (Universität Hamburg, Germany): Using Large-Scale Approaches for Context-Free Testing of Historical Competencies

#### Panel IV: Assessing (Meta)-Conceptual Understanding (Mod.: Christoph Kühberger)

Carla v. Boxtel (University of Amsterdam, Netherlands): Developing Tasks for the Formative Assessment of Historical Causal Reasoning

Lise Kvande (NTNU Trondheim, Norway): Conceptual Understanding and Historical Thinking (Longitudinal Study)

Tagungsbericht *On the Way to an Internationally Shared Assessment of Historical Thinking* (?). 01.07.2016–03.07.2016, Hamburg, in: H-Soz-Kult 11.01.2017.

<sup>10</sup> On this cf. Peter Seixas, Translation and its discontents: Key concepts in English and German history education. *Journal of Curriculum Studies* (2015), pp. 427 – 439; Andreas Körber, Translation and its discontents II: A German Perspective. *Journal of Curriculum Studies* (2016), pp. 440 – 456.