

Poster presentation JURE 2021 conference

Measuring knowledge of historical reasoning and how to teach it in an elementary school PD-programme.

Abstract:

History education researchers emphasize the importance of historical reasoning activities in teaching and learning history (Levstik & Barton, 2015; Levstik & Thornton, 2018; Van Boxtel et al., 2020). Students learn to see history as relevant for their own lives and the developed historical skills as applicable to contemporary problems (Levstik & Barton, 2015).

Historical reasoning is not part of the current elementary school history curriculum in the Netherlands. We developed a two-year professional development programme to prepare grade 3-6 teachers to develop and teach historical reasoning lessons. Teacher professional development is considered to be an essential ingredient in creating and maintaining high quality education (Clarke & Hollingsworth, 2002).

To measure the effectiveness of our programme we developed a series of three paper-based knowledge-skills tests that will measure increasing historical reasoning skills and growth in participants' ability to design historical questions or learning activities to engage students in historical reasoning. We developed and piloted a test to measure teachers' knowledge of historical reasoning and of inquiry-based learning activities. The test contained 21 questions, with 12 multiple-choice questions and 9 short-answer questions. Analysis of the results of the pilot show that $p(\text{average})$ is 0,56. This is within the norm for a combined open and three-option mc-test. Questions which discriminated insufficiently were adapted or removed. Subsequently, experts in the construction of national history tests were asked to evaluate the validity of the ensuing instrument. The labels attached to the questions were checked to increase validity.

The adapted pretest consists of 30 questions: 12 mc-questions and 18 short-answer questions. Due to the adaptations (more items and more items that question the same subconstruct) reliability of the test increased from 0,48 to 0,66. Twenty teachers participated in the pretest. Mean score was 29,7, standard deviation 5,0 and $p(\text{average})$ was 0,66.

Extended summary:

History education researchers have emphasized the importance of historical reasoning activities in teaching and learning history (Levstik & Barton, 2015; Van Boxtel et al., 2020). Historical reasoning activities, such as thinking about cause and consequence, differences and similarities or continuity and change, aim to develop student's ability to formulate and evaluate interpretations of the past by reading, analysing and interpreting historical sources (Van Boxtel et al., 2020). When history lessons aim at doing history instead of knowing history, students develop historical thinking skills and come to understand how historical accounts are created (Levstik & Barton, 2015). These skills are applicable to contemporary problems as well (Levstik & Barton, 2015).

Empirical studies on historical reasoning with children in grade 3 to 6 indicate that, given the chance, students can develop historical reasoning skills on a level fitting their age (Fillpot, 2012; VanSledright & Frankes, 2000). The teacher plays a crucial role in helping students develop these skills (Van Boxtel et al, 2020). But while empirical research has been done on historical reasoning with elementary school students, the question of the preparation and training of elementary school teachers for this task has not been adequately addressed. In studies performed in the United Kingdom and the United States, teachers indicate that they possess inadequate content knowledge to teach the history

curriculum (Temple & Forrest, 2018) and inadequate knowledge about how history is constructed and can be taught (VanSledright & Franks, 2000).

We have developed a two-year professional development programme to prepare teachers of grade 3 to 6 to teach historical reasoning and develop historical reasoning lessons for their students. The aim of professional development is change in teachers' attitudes and beliefs, their classroom practices and in learning outcomes of students. However, professional development programmes have not always proven to be effective in reaching these aims (Clarke & Hollingsworth, 2002).

To measure the effectiveness of our programme, we developed a series of three paper-based knowledge-skills tests that measure historical reasoning skills and growth in teachers' ability to design historical questions. The first test is taken by participants before the start of the programme, the second one halfway and the third one at the end of the programme. The research question is: How can we measure elementary school teachers' knowledge of historical reasoning and skills for suitable inquiry tasks to enhance students' historical reasoning?

Methodology

An instrument of this kind, aiming to measure growth in in-service teachers' knowledge of historical reasoning and how to enhance it, does not yet exist in the Netherlands. There is a national history test for pre-service teachers, but this test includes few questions about historical reasoning and is aimed at students who want to become teachers in secondary schools and are schooled to be history teaching specialists. Our test is aimed at teachers in elementary schools who have had only basic history teaching education.

Based upon theory about historical reasoning and teaching of historical reasoning (Ercikan & Seixas, 2015; Van Boxtel & Van Drie, 2018) we developed a test that consisted of 21 questions, with 12 three-choice questions and 9 short answer questions.

The test was piloted with 22 final-year elementary teachers in training. We investigated the psychometric properties item-difficulty (p) and item-discrimination (R_{it}). In addition, we asked two experts in the construction of national history tests to evaluate the validity of the instrument. Based on the results we adapted the test for use as pretest and developed two parallel posttests.

Preliminary results and findings:

The test consists of 11 questions measuring historical reasoning skills and 10 measuring inquiry skills. In the questions measuring historical reasoning skills the participants receive an historical text or image and a question for which they need to interpret or consider the perspective of the source. The questions measuring inquiry skills present a classroom situation where the teacher has to think of what is needed in a specific phase in the inquiry process, for example a research question that fits that topic or a historical source that can be used.

Mean score was 17,5, standard deviation 3,7 and $p(\text{average})$ was 0,56. $p(\text{average})$ is within the norm for a combined open-question and three-option mc-test. In the analysis of the pilot test, three questions had a p -value between 0,2 and 0,3, indicating that most participants failed. Two questions had a p -value higher than 0,85, indicating easy questions. When assessing these we considered the discrimination factor. We removed one question, adapted the others and added subquestions.

Reliability of the pilot test was low: 0,48. This was influenced by the small amount of questions, the different subconstructs that were measured and the small amount of participants ($N=22$). We added more questions to improve the reliability of the test and an expert checked the questions' labels.

The adapted pretest consists of 30 questions. Due to the adaptations (more items that question the same subconstruct) reliability of the test increased to 0,66. Twenty teachers participated. Mean score was 29,7, standard deviation 5,0 and p(average) was 0,66.

Theoretical and educational significance:

The theoretical significance of this research lies in the advancement of knowledge about how growth can be measured in professional development programmes.

In-service teachers usually never take paper-based tests anymore. Making an accessible test, where teachers show their skills without becoming demotivated, is quite challenging. Having such tests available makes research into teacher professional growth and effects of professional development programmes possible, also in other subject areas.

Literature:

Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and teacher education* 18, 947-967.

Fillpot, E. (2012). Historical thinking in the Third Grade. *The social studies*, 103, 206-217.

Kadriye, E., & Seixas, P. (eds) (2015). *New directions in assessing historical thinking*. New York: Routledge.

Levstik, L., & Barton, K. (2015). *Doing history. Investigating with children in elementary and middle school*. New York: Routledge.

Temple, S., & Forrest, B. (2018). An evaluation of the quality mark for history for primary schools in the UK. *Education 3-13*, 46(6), 661-670.

Van Boxtel, C., Voet, M. & Stoel, G. (2020). Inquiry learning in history. In print.

VanSledright, B., & 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.