The effectiveness of history teacher education concerning the development of lesson planning knowledge: an example from Germany

Jörgen Wolf, 1,* Martin Rothland, 2 Nicola Brauch 1

1 Department of Didactics of History, Faculty of Historical Sciences, Ruhr University Bochum, Bochum, Germany
2 Institute of Educational Science, University of Münster, Münster, Germany
* Correspondence: joergen.wolf@rub.de

Submission date: 21 December 2021; Acceptance date: 26 July 2022; Publication date: 31 August 2022

How to cite

Peer review
This article has been peer-reviewed through the journal’s standard double-anonymous peer review, where both the reviewers and authors are anonymised during review.

Copyright
2022, Jörgen Wolf, Martin Rothland and Nicola Brauch. This is an open-access article distributed under the terms of the Creative Commons Attribution Licence (CC BY) 4.0. https://creativecommons.org/licenses/by/4.0/, which permits unrestricted use, distribution and reproduction in any medium, provided the original authors and source are credited. DOI: https://doi.org/10.14324/HERJ.19.1.08.

Open access
History Education Research Journal is a peer-reviewed open-access journal.

Abstract
Previous studies from German-speaking areas of Europe could not prove any substantial development in the competences and the pedagogical content knowledge of prospective history teachers during teacher training, and thus could not empirically confirm the effectiveness of teacher education. We report on the theoretical framework and the results of a new study on the development of knowledge of lesson planning. In the present study, the development of this kind of knowledge could be confirmed by different test instruments. We found significant differences in the level of knowledge between various cohorts of prospective history teachers (n = 282), but some deficits in history teacher training can also be identified.
Keywords | lesson planning; teacher knowledge; professional knowledge; pedagogical content knowledge; teacher competences; teacher education; teacher training; prospective history teachers

Introduction: modelling and assessing the professional knowledge of history teachers

In contrast to other countries, the effectiveness of teacher training in German-speaking parts of Europe has been examined in recent years primarily according to whether it succeeded in building up professional competences and professional knowledge that are considered relevant (see Tato, 2021; Burroughs et al., 2019; Blömeke, 2017; Park and Suh, 2015). The humanities adapted to this trend, which originated in the mathematical and natural sciences. Therefore, in the past decade, a number of models were presented to describe also the competences and professional knowledge of history teachers (for example, Heuer et al., 2017; Brauch et al., 2015; Husbands, 2011), and test instruments were developed to assess these latent dispositions (for example, Waldis et al., 2019; Resch and Seidenfuß, 2017). Assessing the abilities to plan history lessons also became one focus of empirical research in German history didactics (Litten, 2017; Wolf et al., 2018). (In Germany, Fachdidaktiken [subject-matter didactics] are special academic disciplines of teaching, learning and researching specific school subjects taught in the course of teacher training; see Gundem (2000) and, for history didactics [Geschichtsdidaktik], Rüsen (1987).)

Establishing such instruments opened up the possibility of empirically examining the effectiveness of teacher education with regard to the development of the knowledge considered relevant; however, in comparison to other domains, empirical research in this field in history didactics has only begun with some delay. Research on mathematics teachers, in particular, had already been able to show that professional knowledge, which could be traced back to corresponding formal learning experiences, is built up in a statistically significant and meaningful way in the course of teacher training, thus confirming the general effectiveness of teacher training in this regard (for example, Blömeke et al., 2008; Blömeke and Delaney, 2012; Kunter et al., 2013b; Kaiser et al., 2014). In the meantime, empirical studies on professional knowledge are also available from a large number of other domains. Most of them prove the general effectiveness of teacher education with regard to the development of this knowledge (for example, König and Seifert, 2012; König et al., 2014; Krauss et al., 2017; Lohse-Bossenz and Unger, 2021).

Up to now, the results of the surveys in the field of history didactics could not confirm the effectiveness of teacher education in the same way as the aforementioned studies from other domains. The specific situation in history didactic teacher competence research in the German-speaking area will be briefly introduced below. In several empirical studies with different theoretical assumptions and test instruments, objectively no or only slight increases in subject-specific knowledge of history didactics could be measured in the course of history teacher training (Resch, 2018; Resch et al., 2019; Waldis et al., 2019; Hartmann, 2019). Other investigations using subjective assessment methods revealed that prospective history teachers themselves feel only rather poorly prepared for their later professional activities by their studies (Sauer, 2012; Kanert, 2014). The conclusions of these studies are correspondingly sobering: on the basis of subjective competence assessments, Georg Kanert is of the opinion that the surveyed prospective middle school teachers (n = 259) did not develop a coherent concept of lesson planning, and that the knowledge of subject-specific models and concepts in practice was reduced to ‘buzzword knowledge’ that only legitimises the practicability of the teaching procedure (Kanert, 2014). Ludger Schröer summarises on the basis of interviews and the evaluation of planning documents, as well as vignette testing, that the trainee teachers (Referendare) he studied (n = 16) reveal gaps in didactic vocabulary and discourses (Schröer, 2015). Resch and colleagues could not identify any substantial development of history-didactical knowledge and skills among the prospective secondary school teachers surveyed during their studies (n = 501) or during their traineeship (n = 178) and therefore doubt that the ability to plan and design history lessons develops on the basis of a domain-specific knowledge base acquired during their studies (Resch et al., 2019; Resch, 2018). Carina Hartmann (2019) suspects that the majority of the prospective primary school teachers surveyed (n = 657) would not be able to adequately assess tasks for historical learning, and Monika Waldis et al.
(2019) found that German-speaking Swiss history students (n = 138) focused on the instructional and generic aspects of teaching, but gave less attention to investigations of subject-specific aspects of teaching and pupils’ historical learning. Summarising the research situation to date, there could hardly be any significant developments measured in professional competences during history teacher training in German-speaking parts of Europe, and, in particular, subject-specific (history-didactical) knowledge does not seem to be built up adequately.

Based on an empirical dissertation project, the discourse mentioned above is to be enriched with new empirical evidence, and hence some of the previous assessments are to be put into perspective by the present contribution (Wolf, 2021a). We are therefore interested, first, in whether the effectiveness of history teacher education can be demonstrated by a statistically significant and meaningful increase in knowledge about lesson planning. Furthermore, we investigate to which influencing factors such an increase in knowledge can be attributed. In addition, we examine whether the ascertainable knowledge of prospective history teachers is more subject-unspecific or subject-specific knowledge. Finally, we study whether an interconnection of different types of knowledge related to the tasks of the professional planning activity can be determined, and how the initial concepts of students change on average in the course of teacher training.

First, the theoretical framework and design of the study are presented. Second, the empirical results obtained in this way are summarised. Third and finally, these findings are discussed, whereby suggestions for further research on the competences and knowledge of history teachers are to be made.

Theoretical background

We present two models that allowed us to operationalise the research questions: first, a generic model of planning competence (Figure 1), and, second, a subject-specific model of planning knowledge (Figure 2). Through these models, the relation between planning competence and professional knowledge is clarified and the specific knowledge required for planning history lessons is described. (For a more detailed explanation of the theoretical background, see Wolf, 2021b)

![Figure 1. Planning competence (Source: Wolf, 2021a)](https://doi.org/10.14324/HERJ.19.1.08)
Following dispositional competence approaches, it is assumed in the present model of planning competence that professional knowledge represents the core of teachers' professional competences (Baumert and Kunter, 2013) (see Figure 1).

In addition, by embedding it in an offer-and-use model (Hascher and Hagenauer, 2016), we presume that professional knowledge and affective-motivational aspects of competence are built up or changed through the use of learning opportunities, which can ultimately contribute to an improvement in the quality of teaching and, with appropriate use on the part of the pupils, to greater teaching success.

In contrast to the unspecified models available so far, a specific planning competence in the sense of an independent ‘competence profile’ is postulated here (Oser, 2013), which presupposes specific planning knowledge. We understand planning competence as a continuum from disposition and situation-specific skills to the emergence of observable performance (Blömeke et al., 2015). The planning process itself is modelled as circular or modularised on the basis of the available affective-motivational and, above all, planning knowledge dispositions: individual aspects or subtasks of lesson planning are worked on from different starting points in an unspecified order and with constant references back to each other until a satisfactory solution is found and an overall coherent concept has been developed (Yinger, 1978; Bromme, 1981). Thereby, processes of perception, interpretation and decision making alternate, which have both a creating and a legitimising function for lesson planning (Vogelsang and Riese, 2017; John, 2006).

The model of planning knowledge shows which specific knowledge is considered relevant for the planning of history lessons. We assume that it requires content knowledge, general pedagogical and especially pedagogical content knowledge (history-didactical knowledge) to plan history lessons (Figure 2) (König et al., 2020; Baumert and Kunter, 2013; Shulman, 1987). In this context, content knowledge serves primarily to analyse historical content and media, as well as to identify and provide a historical-scientific or historical-cultural foundation for historical questions (Rüsen, 2005). However, this knowledge also has an effect on the assessment of learning difficulties and, consequently, on methodological aspects and the development of tasks. The general pedagogical knowledge (also general didactic and pedagogical-psychological knowledge) additionally provides interdisciplinary knowledge for structuring lessons, assessing pupils, understanding learning difficulties and ensuring efficient and trouble-free lessons (Guerreiro, 2017; Voss et al., 2015). We presuppose, however, that lesson planning is primarily guided by pedagogical content knowledge (history-didactical knowledge) (Ruszynek and Walton, 2011; Carlson et al., 2019), because this knowledge offers specific information and guidelines for narrowly defined, concrete task areas that have to be dealt with in lesson planning for the subject of history. The superordinate subject-specific and general didactic knowledge, as well as the global concepts and principles of history didactics, only provide background knowledge for these tasks, and for this reason are graphically separated from them (Figure 2).

Based on theoretical assumptions in the history didactics literature, the concrete designing and decision-making level of lesson planning is divided into five different knowledge areas: (1) ‘Goals & Principles’; (2) ‘Phases & Structures’; (3) ‘Media & Methods’; (4) ‘Tasks & Assignments’; and (5) ‘Fit & Adaptivity’ (Wolf, 2021b). The interdependence of these tasks and knowledge areas is made clear in the graphical representation in Figure 2, especially by the cross-section of the knowledge area ‘Fit & Adaptivity’, which encompasses the other areas: all individual decisions in the various task areas of lesson planning must be coordinated with each other to create subject consistency and adaptation to the level of the pupils in order to develop a coherent learning offer for them.

The knowledge areas can be subdivided into further knowledge contents, which are needed for partial tasks of the planning components. In the knowledge area ‘Goals & Principles’, these concrete planning tasks/knowledge contents are constructing themes, defining a historical question as a guideline for the lesson or unit, as well as formulating objectives and promoting competences. In the knowledge area ‘Phases & Structures’, there should be knowledge about the planning context (perspective, unit and lesson planning) and about the phasing of lessons. The knowledge area ‘Media & Methods’ includes the tasks of media selection and presentation, as well as the selection of teaching methods. With regard to ‘Tasks & Assignments’, a distinction can be made between the introduction, elaboration and evaluation of tasks as well as the formulation of learning assignments. In the area of ‘Fit & Adaptivity’, knowledge should be available to diagnose and to adapt planning decisions to the learning requirements of pupils. For all these areas, knowledge elements were identified from the history didactics literature, which are needed for the subject-specific solution of the subtasks (for example, Pandel, 2017; Brauch, 2015; Baumgärtner, 2019; Peters, 2018).
The effectiveness of history teacher education concerning the development of lesson planning knowledge

Design of the study

In order to understand the design of the study, the structure of teacher education in Germany shall be outlined briefly. In Germany, teacher training consists of two phases: a first phase of university training and a subsequent second phase of practical school training, which in North Rhine-Westphalia currently lasts 18 months. During the first phase of training, prospective teachers complete various internships, among which the internship semester (Praxissemester) in North Rhine-Westphalia, which lasts 5 months, takes up the largest share of time during the Master of Education programme (see Terhart, 2019).

To answer the mentioned research questions, a quantitative-descriptive, non-experimental cross-sectional study was conducted, which assessed the lesson planning knowledge of prospective history teachers. Although this assumption is controversial, we assume in this paper that data from cross-sectional studies in cohort designs (often referred to as ‘quasi-longitudinal studies’) can also provide indications of development trends in teacher education (Payne and Payne, 2004; Blossfeld et al., 2009). For this purpose, the sample was tested with the same instruments at one point in time. The sample is divided more precisely into four cohorts of prospective history teachers (n = 282, of which 50.7 per cent female, M = 25.22 years old, SD = 3.45) at different stages of initial teacher training at the Ruhr University Bochum and at five North Rhine-Westphalian centres for practical teacher training (Zentren für schulpraktische Lehrerausbildung):

- 64 first-year students in the Bachelor (BA) programme, who, at the time of the test, had not yet taken any history didactics courses and had no previous practical experience
- 132 Master of Education (MEd) students in the preparatory seminar for the Praxissemester with one previous history didactics course taken
- 52 MEd students in the accompanying seminar for the Praxissemester with at least two history didactics courses and tested after completion of the Praxissemester
- 35 Referendare [trainee teachers], who had already completed their university training and were already at the end of the second (practical) training phase at the time of the test.
In addition to scales on achievement motivation (Rheinberg et al., 2001), teacher self-efficacy (Schwarzer et al., 1999), epistemological beliefs (Maggioni, 2010; Mierwald et al., 2017) and other cognitive (Abitur grade, history grade) as well as non-cognitive characteristics (gender, number of semesters, number of history didactics courses attended, practical school experience measured in days), three test instruments were used to determine the retrievable knowledge of history didactics (that is, pedagogical content knowledge) and of general didactic knowledge, as well as pedagogical-psychological knowledge (that is, pedagogical knowledge):

- the SOSCIE test on history-didactical planning knowledge (Wolf et al., 2018)
- a text-vignette test with an open-ended writing task on lesson planning in school subject history
- a modified form of the test on general pedagogical knowledge used in the Teacher Education and Development Study in Mathematics (TEDS-M) project (König, 2013).

With this triangulative design, it is possible to assess different types of knowledge considered relevant for lesson planning, and to settle their relation to each other. In addition, the vignette testing, with its open-ended task format, offers the possibility of capturing more precisely the relevance system of prospective history teachers concerning the planning aspects of history lessons.

The SOSCIE test captures factual-conceptual and procedural knowledge of history didactics in the above-mentioned areas of knowledge on the basis of 20 closed dichotomous and assignment tasks with a total of 113 items (Wolf et al., 2018). While factual-conceptual knowledge includes knowledge of subject-didactic terms and concepts, procedural knowledge as defined by Anderson et al. (2014) includes knowledge of subject-specific actions (here, for planning history lessons).

The text-vignette test asks the test subjects to give a colleague hints, and to justify which aspects of lesson planning in history are to be considered particularly relevant and how these planning aspects are to be implemented. The texts produced by the test subjects were evaluated according to which content-related areas of knowledge the respective tips could be assigned to. In addition, it was noted whether the corresponding aspects were only (correctly) named (conceptual knowledge), whether their implementation was explained (action knowledge) and whether they were also justified (reasoning knowledge). Furthermore, it was examined whether these elements were named, explained and justified in a subject-specific way or in a general-didactic, respective subject-unspecific way. In addition, under the heading of ‘interconnectedness’, we examined whether subjects were able to both correctly name a knowledge item, explain its implementation, and justify its function for planning. The evaluation was carried out quantitatively, that is, the assignments to the different types of knowledge, areas and contents were counted.

The test on general pedagogical knowledge (PUW-Test) assesses general pedagogical knowledge in the dimensions of dealing with heterogeneity, structuring lessons, classroom management, motivation and performance assessment (König, 2013; König and Blömeke, 2010). For the present study, this test was shortened for reasons of test economy and content, and the evaluation strategy was adapted to that of the SOSCIE test. In the short version we used, the test consists of ten test items, five with a closed response format, and five with an open response format.

Cohort mean values from the knowledge tests were evaluated analytically, and it was first checked whether there was significant difference in average knowledge between the cohorts. In this way, we wanted to find out whether the knowledge of successive cohorts in teacher education differs, assuming increasing knowledge. To find out to which factors the average knowledge differences between the cohorts can be attributed, correlations between knowledge test scores and the other personal trait characteristics of the subjects mentioned above (such as age, gender or epistemological beliefs) were then calculated. Finally, the captured history didactics knowledge was regression-analytically traced back to these cognitive, non-cognitive and affective-motivational initial characteristics to verify the correlations.

Results

The results of both the SOSCIE, the vignette and the PUW tests show that the trainee teachers surveyed have statistically significantly more knowledge at the end of their training than do first-year students and, for the most part, also than the advanced student groups (results of one-way ANOVA or Welch-Test: SOSCIE F(3, 278) = 93.077, p < 0.001, η² = 0.50; Vignette F(3, 256) = 58.777, p < 0.001, η² = 0.36; PUW
The effectiveness of history teacher education concerning the development of lesson planning knowledge

$F(3, 264) = 43.082, p < 0.001, \eta^2 = 0.33$; see Figure 3). The effect sizes are predominantly large with regard to the different mean knowledge levels between the cohorts (see Cohen, 1988). This points to a clear build-up of knowledge in the course of history teacher training (Table 1).

Figure 3. Overall knowledge test results (z-scaled) – the chart shows the deviations from the standardised mean of the sample, the asterisks indicate whether the respective mean cohort value deviates significantly from that of the preceding cohort ($p \leq 0.05$) (Source: Wolf, 2021a)

Regarding the types of knowledge that are queried, there are analogous results determined by both the SOSCIE and the vignette tests, which give more differentiated indications of the development of planning knowledge (Figure 4): the factual-conceptual knowledge assessed by the SOSCIE test, just as the conceptual knowledge assessed by the vignette test, is already built up in an early phase of history teacher education, and it then stagnates at a high level. Here, trainee teachers do not achieve significantly higher values than advanced students in either type of knowledge. In contrast, the procedural knowledge measured by the SOSCIE test, as well as the declarative action knowledge measured by the vignette test, are developed late in the studies and are only substantially built up during the teacher traineeship. Here, the trainee teachers achieve the significantly highest values. This late development is also very clearly recognisable in reasoning knowledge.

With regard to the relationship between the knowledge types of conceptual knowledge, action knowledge and reasoning knowledge, a clear predominance of conceptual knowledge over the other knowledge types is evident. Depending on the cohort, 60.52 per cent to 70.22 per cent of the achieved values can be attributed to conceptual knowledge, whereas action knowledge drops significantly, with 24.44 per cent to 31.32 per cent, and reasoning knowledge is hardly developed in the student cohorts and is only built up substantially in the trainee teachers (3.22 per cent to 9.41 per cent). Accordingly, the levels of interconnectedness are also extremely low: in the case of trainee teachers, only 7.73 per cent of the explained aspects contain both a term and a guideline, as well as the corresponding reasoning; in the case of students, these links can hardly be found in all areas of knowledge. Although statistically there is a difference in knowledge with a strong effect size, which points to an increase in the level of interconnection, especially in the traineeship ($F(3, 256) = 19.979, p < 0.001, \eta^2 = 0.19$), this increasing interconnection of the different knowledge types only takes place at a very low level.
Figure 4. Results knowledge types (z-scaled) – the chart shows the deviations from the standardised mean of the sample, the asterisks indicate whether the respective mean cohort value deviates significantly from that of the preceding cohort \( (p \leq 0.05) \) (Source: Wolf, 2021a)

If we look at the different content-related knowledge areas of planning knowledge, it becomes apparent that the knowledge area ‘Goals & Principles’ is clearly the most developed in terms of effect size, followed by the area ‘Phases & Structures’. In particular, the knowledge contents of constructing themes, formulating objectives and promoting competence, as well as the creation of historical questions, are strongly developed, but also knowledge about unit planning and structuring concepts, history-didactical principles and history-didactical media are developed to an above-average extent relative to the other knowledge contents (see Table 1).

Concerning the phases of teaching, subject-unspecific phase models slightly predominate over subject-specific phase models. Also, a non-subject-specific understanding of teaching methods is evident in all cohorts. For the knowledge areas ‘Tasks & Assignments’ and ‘Fit & Adaptivity’, there are in some cases medium-to-strong effect sizes with regard to the variance of the knowledge characteristics between the cohorts, but, compared with the previously mentioned areas, at a significantly lower level. While the advanced cohorts in the SOSCIE test show quite high solution rates for the test items in the knowledge area of ‘Tasks & Assignments’, explanations of this knowledge area in the vignette texts only take up a marginal amount of space and a significant development cannot be determined here. The questions of how and why teaching should be adapted to the learning requirements of pupils are also rarely described and explained in the vignette texts.
Table 1. Effect sizes of the (single factor) one-way variance analyses (Source: Wolf, 2021a)

<table>
<thead>
<tr>
<th>Total scores</th>
<th>$\eta^2$</th>
<th>$\omega^2$</th>
<th>$d$</th>
<th>Strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOSCIE total</td>
<td>0.50</td>
<td>0.49</td>
<td>2.00</td>
<td>■■■</td>
</tr>
<tr>
<td>Vignette total</td>
<td>0.36</td>
<td>0.36</td>
<td>1.50</td>
<td>■■■</td>
</tr>
<tr>
<td>PUW total</td>
<td>0.33</td>
<td>0.32</td>
<td>1.40</td>
<td>■■■</td>
</tr>
</tbody>
</table>

Knowledge types

<table>
<thead>
<tr>
<th>Knowledge types</th>
<th>$\eta^2$</th>
<th>$\omega^2$</th>
<th>$d$</th>
<th>Strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factual-conceptual knowledge</td>
<td>0.55</td>
<td>0.54</td>
<td>2.21</td>
<td>■■■</td>
</tr>
<tr>
<td>Procedural knowledge</td>
<td>0.25</td>
<td>0.25</td>
<td>1.15</td>
<td>■■</td>
</tr>
<tr>
<td>History didactic knowledge</td>
<td>0.40</td>
<td>0.40</td>
<td>1.63</td>
<td>■■■</td>
</tr>
<tr>
<td>General didactic knowledge</td>
<td>0.08</td>
<td>0.07</td>
<td>0.59</td>
<td>■</td>
</tr>
<tr>
<td>Conceptual knowledge</td>
<td>0.35</td>
<td>0.35</td>
<td>1.47</td>
<td>■■■</td>
</tr>
<tr>
<td>Action knowledge</td>
<td>0.22</td>
<td>0.21</td>
<td>1.06</td>
<td>■</td>
</tr>
<tr>
<td>Justification knowledge</td>
<td>0.16</td>
<td>0.15</td>
<td>0.87</td>
<td>■</td>
</tr>
<tr>
<td>Knowledge interconnection</td>
<td>0.19</td>
<td>0.18</td>
<td>0.97</td>
<td>■■</td>
</tr>
</tbody>
</table>

Knowledge areas

<table>
<thead>
<tr>
<th>Knowledge areas</th>
<th>$\eta^2$</th>
<th>$\omega^2$</th>
<th>$d$</th>
<th>Strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goals &amp; Principles SOSCIE</td>
<td>0.45</td>
<td>0.44</td>
<td>1.81</td>
<td>■■■</td>
</tr>
<tr>
<td>Goals &amp; Principles Vignette</td>
<td>0.43</td>
<td>0.42</td>
<td>1.74</td>
<td>■■■</td>
</tr>
<tr>
<td>Phases &amp; Structures SOSCIE</td>
<td>0.37</td>
<td>0.37</td>
<td>1.53</td>
<td>■■■</td>
</tr>
<tr>
<td>Phases &amp; Structures Vignette</td>
<td>0.11</td>
<td>0.10</td>
<td>0.70</td>
<td>■</td>
</tr>
<tr>
<td>Media &amp; Methods SOSCIE</td>
<td>0.18</td>
<td>0.17</td>
<td>0.94</td>
<td>■</td>
</tr>
<tr>
<td>Media &amp; Methods Vignette</td>
<td>0.07</td>
<td>0.060</td>
<td>0.55</td>
<td>■</td>
</tr>
<tr>
<td>Tasks &amp; Assignments SOSCIE</td>
<td>0.21</td>
<td>0.20</td>
<td>1.03</td>
<td>■</td>
</tr>
<tr>
<td>Tasks &amp; Assignments Vignette</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>Fit &amp; Adaptivity Vignette</td>
<td>0.08</td>
<td>0.07</td>
<td>0.59</td>
<td>■</td>
</tr>
</tbody>
</table>

In order to clarify the research question as to which personal characteristics the knowledge scores can be attributed, the correlative relationships were first determined (see Table 2). The largest highly significant correlations are found for the number of history didactics courses attended. Likewise, the large correlations between number of university semesters as well as age and knowledge indicate that the history-didactical planning knowledge is built up over time in the course of teacher training. The connections to practical experience (measured in days) also appear strong. The Abitur grade as a measure of general cognitive efficiency has a medium, highly significant correlation with test performance. The last grade in the school subject history and the gender of the test subjects appear to be less significant. There are also only isolated and weak correlations between self-efficacy as well as achievement motivation and test performance results. The correlations between the epistemological beliefs and the history-didactical knowledge scores in turn are striking: the more ‘naive’ objectivist beliefs are associated with lower knowledge test scores, while the most elaborate (critical) beliefs are associated with higher knowledge test scores.
Table 2. Correlations of knowledge scores to affective-motivational, cognitive and non-cognitive dispositions (Source: Wolf, 2021a)

<table>
<thead>
<tr>
<th>Variable</th>
<th>SOSCIE total</th>
<th>Vignette (history-didactical)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of history didactics courses</td>
<td>0.610 **</td>
<td>0.557 **</td>
</tr>
<tr>
<td>History semesters</td>
<td>0.595 **</td>
<td>0.417 **</td>
</tr>
<tr>
<td>University semesters</td>
<td>0.566 **</td>
<td>0.396 **</td>
</tr>
<tr>
<td>Age</td>
<td>0.511 **</td>
<td>0.347 **</td>
</tr>
<tr>
<td>Practical experience</td>
<td>0.483 **</td>
<td>0.468 **</td>
</tr>
<tr>
<td>Abitur grade</td>
<td>-0.309 **</td>
<td>-0.273 **</td>
</tr>
<tr>
<td>History grade</td>
<td>0.178 **</td>
<td>0.085</td>
</tr>
<tr>
<td>Gender</td>
<td>0.017</td>
<td>0.145 *</td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>0.049</td>
<td>0.127 *</td>
</tr>
</tbody>
</table>

Achievement motivation
- Presumption of success                      0.190 ** 0.160 **
- Challenge                                   0.031  -0.020
- Interest                                    0.025     0.034
- Fear of failure                             0.005     0.001

Epistemological beliefs
- Objectivist                                 -0.469 ** -0.295 **
- Subjectivist                                0.065     0.042
- Criterialist                                0.344 ** 0.201 **

* Point biserial correlation; ** correlation at the level of 0.01 two-sided significant; * correlation at the level of 0.05 two-sided significant; interpretation, $\rho$ according to Cohen (1988), $>0.10$ small correlation, $>0.30$ medium correlation, $>0.50$ large correlation.

The relationship between the dependent variables of the total score of the SOSCIE test and the history-didactical knowledge score in the vignette test, and the independent variables of affective-motivational, cognitive and non-cognitive personal characteristics listed above, was calculated by two linear multiple regression analyses. It was assumed that all of the predictor variables mentioned above (can) have an equal influence on history-didactical planning knowledge; a moderation/mediation hypothesis was not set up. Both regression models show a high goodness of fit with high statistical significance (Model a: dependent variable SOSCIE total score, $F(14, 244) = 24.082$, $p < 0.001$, corr. $R^2 = 0.556$; Model b: dependent variable history-didactical proportions of vignette score, $F(14, 222) = 11.096$, $p < 0.001$, corr. $R^2 = 0.375$). In a direct comparison, only a few predictor variables show a significant relation to the history-didactical planning knowledge expended (see Figure 5). In both models, the number of history didactics courses attended proves to be a highly significant and relationally strongest predictor. Clear correlations can also be seen with epistemological beliefs; with regard to the Abitur grade, the connection is moderate but there are only marginal correlations between lesson planning knowledge and practical experience.
Discussion and outlook

The study presented here is subject to a number of limitations, which restrict its significance and scope. On the one hand, it is a non-representative random sample. The most representative sample can be claimed for the tested students of the Ruhr University Bochum, where about 70 per cent of the training cohorts of one year were recorded. However, it was also not possible to achieve global representativeness in the other empirical studies on history didactics described above. It is still a research objective to realise history didactics studies for the objective recording of teacher competences or professional knowledge of such quality in Germany. The same can be said with regard to the research design and the chosen statistical evaluation procedures. Our data and the assumption of developmental trajectories are based on a cross-sectional study in a cohort design. Here, test results of different individuals are compared with each other via their cohort affiliation. However, for the valid identification of intra-individual developments, test results of the same individuals would have to be compared in a longitudinal design, at best using appropriate statistical procedures such as multilevel and latent change models within the framework of item response theory (De Ayala, 2013). However, due to the difficulties of recruitment and panel mortality in longitudinal studies, other large-scale studies on professional knowledge also chose a design and procedures comparable with those of our study in order to obtain initial evidence on the effectiveness of teacher education, and it turns out that very similar results are obtained through both study designs, and also when different test theories are applied. Finally, the training aspects in the present study were recorded under purely quantitative aspects and in the very rough categories of participation in history didactics courses and practical experiences measured in days, and not further differentiated. This produces a purely quantifying picture of formal learning opportunities. Differences in quality, that is, what is offered and how it is used, cannot be recorded in this way, and the effects of teacher education can accordingly hardly be traced back finely enough to such conditions. Therefore, indications for a reorganisation of history didactics teaching or subject-specific teacher training can hardly be formulated, or only very vaguely.

Despite these limitations, the results of our study should prove interesting for the discourse on history didactics, and for research on lesson planning competence in general, as they provide strong evidence for a general positive effectiveness of history teacher education related to the development of history-didactical knowledge. In contrast to the studies by Resch and colleagues (Resch, 2018; Resch et al., 2019), as well as by Hartmann (2019), at least significant medium-to-strong effect sizes were found relating to the different levels of knowledge between prospective history teachers in different training phases. These effect sizes correspond to the values from studies in other domains, and in some cases
The effectiveness of history teacher education concerning the development of lesson planning knowledge

even clearly exceed them (see, for example, Blömeke et al., 2008, 2010; Kunter et al., 2013a; König and Seifert, 2012; Krauss et al., 2017).

Above all, the knowledge area ‘Goals & Principles’ is strongly developed, which, according to theory and empiricism of history didactics, comprises the most essential and most specific parts of lesson planning in history (Pandel, 2017; Litten, 2017). Within the framework of teacher training, prospective history teachers seem to build up and internalise those disciplinary terms and concepts which, from a theoretical perspective, should support subject-specific lesson planning, or at least enable communication about lesson planning.

What could nevertheless be criticised according to our data, is the depth and interconnectedness of the knowledge examined. In particular, the results of the vignette testing – in view of the strong preponderance of conceptual knowledge and low interconnection levels – show that prospective history teachers can name relevant planning contents in detail, but less knowledge can be demonstrated about how these components are implemented and why such planning elements are considered relevant. It can be proven that history-didactical terms and concepts are already built up early in the studies and, measured by the solution rates of the SOSCIE test, are also predominantly handled confidently. But, in view of the significantly lower level of procedural and declarative action knowledge on lesson planning, as well as the barely developed knowledge of reasoning, it can still be asked to what extent prospective history teachers already (can) become sufficiently competent through the two phases of teacher training in Germany.

However, this question can hardly be answered conclusively for various reasons. On the one hand, to date there are no precise subject-specific standards against which empirically determined training results can be measured (König et al., 2019). On the other hand, the test items and the knowledge content enquired about are based on normative specifications derived from history didactics literature. Thus, in our tests, as well as in the other history-didactical knowledge and competence tests mentioned above, it is predominantly (academically produced and abstracting) theoretical-formal knowledge that is measured, and less practical knowledge (gained from practice and related to concrete cases) (see Fenstermacher, 1994). It is thus understandable that such defined planning knowledge can be empirically attributed primarily to history didactics courses, and that practical experience seems hardly to influence such knowledge. However, it is disputed whether or to what extent knowledge, and especially such theoretical-normative knowledge, actually guides concrete action practice (Grima-Farrell et al., 2019; Neuweg, 2002). Empirically, this could be clarified by determining the prognostic or action validity of knowledge tests designed in this way, for which the test results must be related to empirically observed (planning) performances. Such validation studies have also only recently been carried out in other domains, and they have not yet been completed (see, for example, Blömeke et al., 2022; Kaiser and König, 2020). Observations of concrete planning practice could provide additional insights into which operations and heuristics history teachers really use in planning. In this way, the motivation for a revision of the previous theoretical-normative assumptions on lesson planning in history could be expected (Rothland, 2021; König and Rothland, 2022).

Declarations and conflicts of interest

Research ethics statement

The authors conducted the research reported in this article in accordance with the ethical guidelines defined by Deutsche Forschungsgemeinschaft (DFG).

Consent for publication statement

The authors declare that research participants’ informed consent to publication of findings – including photos, videos and any personal or identifiable information – was secured prior to publication.

Conflicts of interest statement

The authors declare no conflicts of interest with this work. All efforts to sufficiently anonymise the authors during peer review of this article have been made. The authors declare no further conflicts with this article.
The effectiveness of history teacher education concerning the development of lesson planning knowledge

Notes

1Abitur is a qualification granted at the end of secondary education in Germany. As a matriculation examination, Abitur can be compared to A levels, the Matura or the International Baccalaureate Diploma, which are all ranked as Level 4 in the European Qualifications Framework.

References


Heuer, C., Resch, M. and Seidenfuß, M. (2017) ‘“What do I have to know to teach history well?” Knowledge and expertise in history teaching – a proposal’. Yesterday and Today, 18 (1), 27–41. [CrossRef]


The effectiveness of history teacher education concerning the development of lesson planning knowledge


