

# The untapped potential of mixed-methods research approaches for German history education research

Manuel Köster\* and Holger Thünemann – *University of Cologne, Germany*

## Abstract

Despite some pioneering studies, mixed-methods research approaches are uncommon in the German history education community, in contrast to the general increase in mixed-methods research in the educational and social sciences. Conversely, German history education research currently appears to favour quantitative methods as opposed to qualitative approaches – at least in larger research projects. In this paper, we argue for a more inclusive research approach combining qualitative and quantitative methods. Discussion of a pioneering study from the 1980s (Jeismann *et al.*, 1987) highlights implementation of this unusual approach to history education research in Germany. To illuminate the added value of such a mixed-methods research approach, we discuss two published German studies that respectively rely on quantitative (Trautwein *et al.*, 2017) and qualitative (Köster, 2013) research methods. A mixed-methods approach might have illuminated each study's 'blind spots'.

**Keywords:** history education research; mixed methods; historical thinking; large-scale assessment; reading comprehension

## Introduction

While uncommon in German history education research, increasingly social science and education research studies combine qualitative and quantitative approaches, as documented in handbooks since 2000 (for example, Creswell, 2003; Creswell and Plano Clark, 2007; Teddlie and Tashakkori, 2009; Gläser-Zikuda *et al.*, 2012; Kuckartz, 2014), the *Journal of Mixed Methods Research* founded in 2007, and the Mixed Methods International Research Association (MMIRA). Mixed-methods research approaches combine a range of research methods, procedures and techniques from different methodological areas. This paper contains several highly contrasting mixed-methods research models in a brief overview of different ways of combining qualitative and quantitative approaches that focuses on the benefits of mixed methods research for history education. Accordingly, the paper describes a pioneering exemplar study that utilized a mixed-methods approach in contrast to two other studies – one mainly qualitative, one quantitative – that could have benefited from a mixed-methods approach. Our aim is to demonstrate the potential of a mixed-methods approach, largely unrealized, and not to criticize the single-method approach. Some of the arguments made in this paper are already part of a German-language publication (Prinz and Thünemann, 2016).

Due to different research traditions and their corresponding paradigms, researchers working with quantitative methods frequently display a certain scepticism

towards the methodological foundations, the data collection and interpretative methods employed in qualitative research – and vice versa (Kelle, 2014: 154–7). While quantitative researchers may consider that qualitative research produces little of substantive value, in turn qualitative researchers may consider the quantitative paradigm as of limited value. Lamnek (2010: 16) argues that the two research approaches have ‘differing views on the nature, the structure, the character, the essence of reality’. Since 2005, efforts to overcome this divide in Germany (Mayring, 2012) and the USA (AERA, 2006) developed criteria applicable to both qualitative and quantitative research.

Kelle (2008: 14) sees this divide as an indicator of a ‘paradigm war’, where, based on scientific and epistemological assumptions, different research problems and their questions demand particular research procedures, plans and methodologies. Kelle (ibid.: 53) argues that ‘inflexibly allotting methods to certain epistemological positions or “paradigms” [was] by no means imperative’. Instead, he points out that research practice proved that qualitative and quantitative methods were not at all incompatible. Combined productively, they can strengthen and mutually complement each other, and reinforce the validity and viability of mixed-methods models in educational research from asking valid questions onwards.

According to Tashakkori and Teddlie (2003: 711), mixed-methods research can be defined as ‘a type of research design in which QUAL and QUAN approaches are used in types of questions, research methods, data collection and analysis procedures, or in inferences’. This broad definition allows for a considerable level of heterogeneity, and enables researchers to make decisions in a number of fields, depending on the goals of the research project. Creswell and colleagues (2003: 215–23) list four criteria that need to be considered when deciding on an appropriate mixed-methods design:

1. Implementation: in which order do qualitative and quantitative data collection take place?
2. Priority: which methodological approach is given priority?
3. Integration: at which step of the research process are qualitative and quantitative data or results integrated?
4. Theoretical perspectives: which role does the theoretical perspective play? Is it implicit or does a theory frame the whole research design?

These criteria outline the dimensions that make up a mixed-methods research design. There is a range of variables:

- four different ways to combine qualitative and quantitative data
- different times for integrating additional methods
- different points for analysing data during data interpretation
- the role of theoretical perspectives, from them being implicit, to having an explicitly dominant role in shaping the research.

The design types listed by Creswell and colleagues allow for a total of 72 such combinations, from which Creswell (2003) derives 6 main designs.

## Sequential designs

Three of Creswell’s (2003) designs are sequential, which means that two separate studies using different methodological approaches follow each other, with the results of the first study informing the next. Two different types of sequential design are possible: depending on whether the project begins with a qualitative or a quantitative study, it is an ‘explanatory design’ or an ‘exploratory design’.

An explanatory design is a sequential design where quantitative methods are used for data collection and interpretation first. The qualitative study that follows is intended to allow for a deeper understanding of statistical computations and models, and to explain unexpected or otherwise inexplicable results obtained through quantitative analyses (Creswell, 2003: 215; Creswell *et al.*, 2003: 223, 227; Creswell and Plano Clark, 2011: 81–6). Exploratory designs, on the other hand, are the reverse side of explanatory sequential designs. They begin with a qualitative study, whose findings are then substantiated, tested and generalized by using quantitative methods (Creswell, 2003: 215–16; Creswell *et al.*, 2003: 227–8; Creswell and Plano Clark, 2011: 86–90). Following Mayring (2001), Kuckartz (2014: 67) argues that the term *verallgemeinerndes* or *generalisierendes Design* (both can be translated as ‘generalizing design’) is more fitting than ‘exploratory design’. In both explanatory and exploratory design, the focus can be either on the qualitative or on the quantitative phase, or both can be equally important. Finally, ‘transformative designs’ are characterized by a shared theoretical perspective underlying both phases of the investigation; while there is a sequential order to both phases, neither of them is given priority (Creswell and Plano Clark, 2011: 96–100). The two investigations can be transferred into the other study format at the end of the project. They can be connected from the outset or one can be integrated into the other. In any case, the project is framed by a transformative theoretical perspective that informs all methodological choices (Kuckartz, 2014: 67).

## Parallel designs

The other three of Creswell’s (2003) main design types are parallel, non-sequential qualitative and quantitative studies on the same subject. Implementation of these enables a fuller, more holistic understanding of a common research problem through comparing their findings and conclusions. Often, research reports are created independently, and only later are results related to one another (Teddlie and Tashakkori, 2009: 152) in order to generate a so-called meta-inference, ‘a conclusion generated through an integration of the inferences that have been obtained from the results of the QUAL and the QUAN strands of an MM [Mixed Methods] study’ (*ibid.*). Among the types of parallel design, there is ‘triangulation design’, whose goal is ‘to obtain different but complementary data on the same topic’ (Morse, 1991: 122), usually ‘in an attempt to confirm, cross-validate, or corroborate findings within a single study (Creswell *et al.*, 2003: 229). Ideally, both the qualitative and quantitative parts of the study should be given equal weight. However, as Creswell and colleagues (*ibid.*) note, in reality, one of the methods is frequently prioritized. This model is often called ‘concurrent triangulation’. It is equivalent to Denzin’s (1970) classic ‘between-method triangulation’, where independent quantitative and qualitative enquiries are conducted and compared (Kuckartz, 2014: 67). Creswell and Plano Clark (2011) call this strategy a ‘convergent design’, since parallel lines of research are analysed independent of each other and then related to one another.

A strength of this design type is its efficiency, since qualitative and quantitative methods can be applied simultaneously. This also means that a single research team initially interprets the data, thus avoiding possible confusion from the involvement of more than one group (Creswell and Plano Clark, 2007). However, implementing triangulation design is often viewed as challenging (Creswell and Plano Clark, 2011: 79) because concurrently implementing two different research paradigms and their methodologies requires experienced, skilled researchers, and possibly the formation of research groups whose members have complementary skills. Nevertheless,

triangulation is among the four most popular mixed methodologies, alongside embedded, explanatory and exploratory.

A second parallel design is the 'concurrent transformative strategy', which is characterized by its underlying theoretical perspective, similar to the sequential transformative design discussed above. This perspective can draw from different schools of thought, such as critical theory or participatory research (Creswell, 2003: 26), and informs either the overarching research question or goal, or both.

While both paradigms have equal weight in the transformative strategy, one of the two methodological approaches dominates the 'concurrent nested strategy', also called 'embedded design'. It uses the secondary, complementary method to generate sub-questions to complement the dominant methods that help shape the early stages of data collection. A common way to do this is to embed a qualitative study (often already published) within a larger quantitative study, and to consider qualitative findings alongside the quantitative (Creswell and Plano Clark, 2011: 92–3). Within the research project, using the embedded method, the research team can use qualitative methods to investigate a particular aspect of the research field outside the purview of the dominant quantitative method but relevant to it.

A strength of this strategy is its economy, both in terms of effort and in terms of requiring fewer kinds of data. Additionally, this design type is particularly suited to educational contexts, as its focus is on quantitative methods needed for experimental designs or correlational analyses (ibid.: 94). However, Creswell and Plano Clark (ibid.: 94–5) also point out numerous drawbacks of this research paradigm. First, the researchers must specify why the qualitative data is necessary to complement and enrich the larger quantitative study. Second, integrating the results of both studies can be challenging, since different methods are being used in order to answer different research questions. Unlike triangulation designs, embedded designs do not strive to relate different data sets to one another in order to answer a shared research question. Rather, results are intended to be published independently and in different publications.

## Other designs

According to Kuckartz (2014: 87–90), 'transfer designs' are characterized by the fact that data types are transferred into each other, which is followed by a focused integrative analysis that utilizes only this transferred data type. There are two types of transfer design, intended for either 'quantitizing' or 'qualitizing' data. Again, according to Kuckartz (ibid.), quantitizing data is the more common approach. While this means that qualitative data is transferred into countable units, qualitizing means that quantitative data is categorized or transferred into a verbal statement, in order to produce a more holistic and integral case study (Bazaley, 2009). Whether transfer designs truly are a type of mixed-methods research has been called into question. For example, Burzan (2015: 5) states that 'a quantitative analysis of non-standardized data does not seem to follow the principles of qualitative research, and qualitative research based on standardized data appears hardly feasible'. (All translations from German-language publications were made by the authors of this paper.)

In addition to the designs discussed so far, more complex designs consisting of more than two phases are occasionally applied. This can be achieved either by combining two two-step designs into a three-step design (Kuckartz, 2014: 91), usually with the middle step taking precedence, or through 'multiphase' (Creswell and Plano Clark, 2011: 100–4) or 'fully integrated multi-strand designs' (Teddlie and Tashakkori, 2009: 156–60), where qualitative and quantitative methods are combined and often transferred into one another in a dynamic and interactive fashion.

## Existing mixed-methods research on history education

Even though the typology of mixed-methods research discussed above has not been systematically considered in German research on history education, similar approaches have played a role since the 1970s – albeit a relatively minor one. In the following paragraphs, we discuss an early study utilizing a mixed-methods design (Jeismann *et al.*, 1987). Like many empirical studies in Germany, it draws upon the principles of historical thinking and the category of historical consciousness developed mainly by Karl-Ernst Jeismann, Jörn Rüsen, Bodo von Borries and Hans-Jürgen Pandel (see Seixas, 2016; Schönemann, 2017; Kölbl, 2017). The research project was initiated in the 1970s, and was published in 1987. The authors – history education researchers with a background in history, as well as psychologists – investigated the effect that the teaching unit ‘The separation of Germany and the formation of two German states’ had ‘on the historical consciousness of adolescents’ (Jeismann *et al.*, 1987: 11). (Another early example combining qualitative and quantitative methods in a creative and convincing manner is Von Borries *et al.*, 2005.)

Methodologically speaking, this was a quasi-experimental intervention design consisting of two preliminary studies and one main investigation (Jeismann *et al.*, 1987: 33) intended to examine an assumed causal relationship (*ibid.*: 71). Aspects of historical consciousness and historical thinking addressed were factual knowledge (corresponding to Jeismann’s ‘factual analysis’, *Sachanalyse*), factual judgements (*Sachurteil*) and value judgements (*Werturteil*). As in any intervention, a particular difficulty was to allow for the influence of potential variables and to attribute any changes in the complexly inter-related variables to the intervention (*ibid.*: 77).

The sample consisted of ‘thirty year nine classes, consisting of 653 [15-year-old] students’ drawn from the highest and the lowest tiers of the three-tiered German educational system, which were then randomly allotted to the treatment or control condition and tested in a pre-post-test design (Jeismann *et al.*, 1987: 79). In all phases of the project, open-item questionnaires were used (*ibid.*: 33–6). These were aimed at the aspects of historical thinking mentioned above, and designed in such a way that students received a number of points for each answer or partial answer they provided. Thus, ‘for each characteristic a distinctive number of points could be assigned’ (*ibid.*: 80; see also *ibid.*: 39–41, 46–9). (On the partial credits system used by Jeismann *et al.* in 1987, now see VanSledright, 2014: 58–66. He, however, uses weighted multiple-choice items, not open items.) In order to assess students’ political views and beliefs, Likert-scaled closed items were used. In addition to these steps, the authors conducted subsequent teacher interviews (Jeismann *et al.*, 1987: 99ff.).

In relation to the taxonomy outlined above, the Jeismann *et al.* (1987) research project was a complex multiphase design consisting of two pre-studies, a main study with a pre- and a post-test and, finally, teacher interviews. Since qualitative data from open items were quantitized – a main challenge of the project – a transfer design was included in the multiphase design. Jeismann and his fellow researchers concluded that the intervention teaching unit incorporating Jeismann’s concept of historical consciousness led to ‘an increase in factual knowledge and factual judgements’, while ‘results were inconclusive as towards historical value judgements’ (*ibid.*: 106; see also *ibid.*: 194–7). Apparently, the fact that results concerning students’ value judgements were unclear was in no small part due to a relatively ‘large number of reasoning tasks whose answers were either non-evaluable ... or, in particular, that were not answered at all’ (*ibid.*: 106; see also *ibid.*: 195). In other words, there were numerous values missing from the evaluation.

## A plea for increased mixed-methods approaches in history education research

Although empirical studies combining qualitative and quantitative methodology date back to the 1980s, mixed-methods research designs are not particularly common in the German research community on history education. As a relative latecomer to systematic empirical research, the didactics of history was not very much affected by the 'science wars' (Ross, 1996) of the 1980s and 1990s. While the pendulum of academic focus tended to swing between qualitative and quantitative approaches, history education research in the late twentieth century, as well as in the 2000s, was marked by a certain methodological experimentalism, as well as by increased methodological awareness (Bracke *et al.*, 2014). This situation appears to have changed somewhat over the last few years, as history education researchers in Germany increasingly seem to adhere to one research paradigm exclusively.

Particularly in the context of larger projects, qualitative approaches appear to be becoming increasingly rare. Instead, researchers seem to be inspired mainly by educationalists and psychologists who use quantitative methods in order to accurately 'measure' student thinking. Outside of PhD projects, qualitative approaches currently seem somewhat relegated to the sidelines. This trend is visible in the volumes documenting the empirically focused, biennial conference *geschichtsdidaktik empirisch*, which increasingly consists of studies using a quantitative approach (compare Ziegler and Hodel, (2009) documenting the first conference, and Waldis and Ziegler (2017), which contains papers presented at the most recent conference). In line with similar developments in neighbouring fields and in educational policy, current notions of scientificity seem to be heavily informed by methodological conventions and epistemological concepts of quantitative research. For example, a focus group in the KGD, the association of German history education researchers, now aims to set empirical standards to which any paper published in the association's journal is supposed to adhere. The standards discussed thus far lean heavily towards quantitative methods and a positivist research paradigm. Similarly, research programmes funded by the German ministry of research and education are geared towards quantitatively focused projects. If such a view of the nature of empirical research were to be exclusive, this would be an unfortunate development, since both schools of research clearly provide specific sets of affordances and constraints, of bright and blind spots. As we have outlined above, a combination of qualitative and quantitative methods can contribute to illuminating such blind spots by sensitizing researchers to the potential, as well as to the limitations, that arise from the specific epistemological position underlying both methodological paradigms (see Köster, 2016: 11–14, 30–2).

In the final section of this paper, we intend to further illuminate the potential of mixed-methods approaches – epistemological differences between qualitative and quantitative schools of thought notwithstanding – by re-analysing two existing studies that rely heavily on either qualitative or quantitative methods, and by discussing how they could have benefited from a mixed-methods approach.

### Adding qualitative elements to a mainly quantitative study

In the HiTCH-project (Historical Thinking – Competencies in History; Trautwein *et al.*, 2017), German, Austrian and Swiss empirical researchers with a background either in history or in educational research cooperated in order to try and 'develop a test that



can measure students' historical thinking competences in large-scale assessments' (ibid.: 11). The HiTCH test was based on the FUER model of historical thinking (Körber *et al.*, 2007), which has been discussed intensively in the academic community (Barricelli *et al.*, 2012: 219–21; Thünemann, 2016: 37–43, 46–8). The FUER-Geschichtsbewusstsein project (Förderung und Entwicklung eines reflektierten Geschichtsbewusstseins; or Promotion and Development of a Reflected Historical Consciousness) was an EU project from which a competence model of historical thought emerged. This FUER model lays out four areas of historical competency: Competencies in Inquiring (historically), Competencies in Applying Historical Methods, Competencies in Historical Orientation and Historical Subject Matter Competencies (translation and capitalization taken from Körber, 2015: 40).

The main investigation involved 2,853 Year 9 students (median age: 14.41 years) from different schools, school types, countries and states. The main investigation was intended to measure students' competencies 'objectively, reliably, validly and using standardized test items' (Trautwein *et al.*, 2017: 56, 82). The test did not address all levels of the FUER model, but only its 'intermediary, conventional level of competence' (ibid.: 122), since 'qualitative leaps' to an elaborate, trans-conventional level 'cannot be measured' with the HiTCH test (ibid.). In its final form, the test consists of 91 items (the item pool originally consisted of 106 items) in 15 task sets. While 8 task sets address students' methodological competencies (re- and de-constructing history) and 4 sets test for their subject matter competencies, there is only 1 task set assessing competencies in historical orientation and 2 addressing students' enquiry competencies (ibid.: 89–91). Presumably, this 'slight imbalance' (ibid.: 119) is in no small part due to the fact that the latter competencies, while absolutely integral to historical thinking, are very hard to assess with standardized, closed items (ibid.: 118–19).

According to the authors, while an 'empirical "confirmation"' of the FUER model in its four dimensions through the HiTCH study was neither possible nor intended, it enabled 'the successful construction of a large-scale test of historical thinking as a whole' (ibid.: 119). This, they conclude, was 'an important step', since it constituted 'overcoming the obstacles connected with psychometrically measuring a complex scholarly discipline such as history' (ibid.: 128). However, the almost exclusive focus on quantifiable results bears certain limitations, which the authors try to reflect on themselves (ibid.: 116–23). These could have been mostly compensated for by a mixed-methods design.

If the design of the pioneering HiTCH study was to be expanded by open tasks where students had to pose historical questions and to create historical narratives, this could benefit the assessment and development of historical competencies in at least three ways. First, this would mean that not only an intermediary, conventional level of historical thinking could be addressed, but also an elaborate, trans-conventional one. This is especially important since it is at this level that the generative or creative character of problem-solving – a defining property of competencies (Pandel, 2005: 24) – comes into play. Second, this would allow for an appropriate inclusion of historical enquiry and orientation, two fundamental dimensions of historical thinking under-represented in the current task sets. This is especially important since it would enable researchers 'to address individual [historical] orientation' (Trautwein *et al.*, 2017: 74) and the interplay of 'the interpretation of the past, the experience of the present and expectations towards the future' (Jeismann, 1997: 42). The creators of the HiTCH test decided to 'deliberately abstain' (Trautwein *et al.*, 2017: 74) from this. Third, using open items and qualitative interviews can provide insights into the specific challenges that students face when turning their historical competencies into competent performance.

The didactics of history, as ‘the science of historical learning’ (Rüsen, 1989: 84), could particularly benefit from this last aspect, since it would contribute to diagnosing and fostering competencies (Thünemann, 2016: 46–8). For a first glimpse of such challenges, see Werner and Schreiber (2015).

## **Adding quantitative elements to a mainly qualitative study**

Reassessing published research projects can sometimes seem like criticizing someone else’s work with the benefit of hindsight. We therefore decided to further illustrate the potential of mixed-methods research by re-evaluating a study conducted by one of the authors of this paper. In his doctoral dissertation (PhD thesis), Manuel Köster (2013) investigated how students’ reading comprehension when engaging with two different texts on the Holocaust was influenced by their prior beliefs and value judgements (Jeismann, 2000: 63–9). Specifically, Köster wanted to know whether students would be able to recognize that the two texts they were asked to read put forth diametrically opposing views on the knowledge of, and culpability for, the Holocaust attributable to the non-Jewish German population during National Socialism.

Using a qualitative approach, Köster was indeed able to trace the influences of students’ preconceptions on their reading comprehension. This was especially true for those learners who strongly identified with Germany. These students tended to adapt the situation model (Kintsch and Van Dijk, 1978) for both texts to their pre-existing beliefs by focusing only on select aspects. They displayed four distinct strategies in doing so (Köster, 2013: 104–5), frequently without recognizing that both texts provided different interpretations of the same event. Interestingly, students who were either born outside Germany or were (first- or second-generation) descendants of immigrants were less inclined to employ these strategies (*ibid.*: 213–22).

This study already employed a mixed-methods design, inasmuch as it combined a quantitative survey ( $n=272$ ) with qualitative interviews ( $n=50$ ) (*ibid.*: 52–62). While the survey was intended to document students’ preconceptions about, and knowledge of, National Socialism and the Holocaust, as well as a number of other factors, to bring to light possible correlations between these factors and to select participants for the interviews based on this data, the main focus of the study was on the focused interviews (Merton and Kendall, 1946). The design was thus somewhat similar to an explanatory design, as it began with a quantitative investigation, the results of which informed the qualitative phase of the project. However, the quantitative survey did not test students’ reading comprehension. Köster’s PhD thesis therefore did not fulfil the main purpose of an explanatory design, that is, using qualitative methods to further illuminate findings of a quantitative investigation.

In order to benefit from the advantages of an explanatory design, Köster would have had to measure students’ reading comprehension in the quantitative stage of the project too. An ideal – if considerably more time- and labour-intensive – solution would have been to combine the existing survey with both a standardized test of reading ability and a survey investigating reading comprehension of the texts used in the qualitative stage. A standardized test of reading ability would have provided a measure of students’ reading ability independent of the factors whose influence was being investigated in the project. Such a measure could have helped prevent the attribution of too much influence to preconceptions and student identification with Germany. Without such a measure, some of the observed phenomena might have been misconstrued as stemming from domain-specific factors, rather than from general



reading ability – even though the perceived influence of preconceptions and other domain-specific factors was of course weighed against other factors while interpreting the interview data.

A survey investigating reading comprehension of the texts used in the qualitative stage, on the other hand, would have provided a much clearer picture of how much such domain-specific factors influence reading comprehension in a wider sample of participants. As it is, participants were selected based on their performance in a questionnaire that surveyed specific factors that were assumed to influence their comprehension of the two contrasting texts. Thus, a quantitative test employing multiple- or single-choice items to document student comprehension of these texts could not only have brought to light the possible influence of other factors not accounted for in the original questionnaire (which could then also have been used as additional selection criteria for the interviews), but could also have shed light on whether the four strategies observed in the (original) qualitative phase of the project are common among students in general or if they (positively or negatively) correlate with certain preconceptions. While every effort was undertaken to prevent this from happening, the danger of viewing the data with ‘theoretical blinkers’, of somewhat impressionistic interpretations influenced by theoretical assumptions, is always prevalent in qualitative data. An additional quantitative element could have helped to further prevent this from happening and to substantiate the findings.

## Conclusion

As outlined above, the potential of mixed-methods research is not widely utilized in the German research community on history education – despite some pioneering studies. Possible explanations range from epistemological beliefs and positions, through a (perceived) lack of methodological skills and knowledge, to the fact that empirical studies in the relatively small community of history education research are often conducted in PhD projects and other contexts of very limited resources. Nevertheless, combining qualitative and quantitative perspectives on a research object can certainly lead to a more holistic view of the problem under investigation.

## Notes on the contributors

**Dr Manuel Köster** is a tenured senior lecturer (*Akademischer Oberrat*) at the University of Cologne, Germany. With a background in history, he has published in German and English on the interconnection of history and language, on empirical history education research, on public history and on the theory of history education. Publications to be released in 2019 include an international edited volume on history education research and a book on learning and assessment tasks.

**Dr Holger Thünemann** is Professor for the Didactics of History at the University of Cologne, Germany. He previously worked as a school teacher and at the universities of Münster and Freiburg. He has published numerous German- and English-language books, edited volumes and articles on public history, on history textbooks and on empirical research.

## References

AERA (American Educational Research Association) (2006) ‘Standards for reporting on empirical social science research in AERA publications’. *Educational Researcher*, 35 (6), 33–40.

- Barricelli, M., Gautschi, P. and Körber, A. (2012) 'Historische Kompetenzen und Kompetenzmodelle'. In Barricelli, M. and Lücke, M. (eds) *Handbuch Praxis des Geschichtsunterrichts* (Vol. 1). Schwalbach am Taunus: Wochenschau Verlag, 207–35.
- Bazaley, P. (2009) 'Computer-Assisted Integration of Mixed Methods Data Sources and Analyses'. In Tashakkori, A. and Teddlie, C. (eds) *Foundations of Mixed Methods Research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Thousand Oaks: Sage, 431–476.
- Bracke, S., Flaving, C., Köster, M. and Zülsdorf-Kersting, M. (2014) 'History education research in Germany: Empirical attempts at mapping historical thinking and learning'. In Köster, M., Thünemann, H. and Zülsdorf-Kersting, M. (eds) *Researching History Education: International perspectives and disciplinary traditions*. Schwalbach am Taunus: Wochenschau Verlag, 9–55.
- Burzan, N. (2015) 'Review: Udo Kuckartz (2014) *Mixed Methods: Methodologie, Forschungsdesigns und Analyseverfahren*'. *Forum: Qualitative Sozialforschung*, 16 (1), Article 16. Online. <http://nbn-resolving.de/urn:nbn:de:0114-fqs1501160> (accessed 30 May 2018).
- Creswell, J.W. (2003) *Research Design: Qualitative, quantitative, and mixed methods approaches*. 2nd ed. Thousand Oaks, CA: SAGE Publications.
- Creswell, J.W. and Plano Clark, V.L. (2007) *Designing and Conducting Mixed Methods Research*. Thousand Oaks, CA: SAGE Publications.
- Creswell, J.W. and Plano Clark, V.L. (2011) *Designing and Conducting Mixed Methods Research*. 2nd ed. Thousand Oaks, CA: SAGE Publications.
- Creswell, J.W., Plano Clark, V.L., Gutmann, M.L. and Hanson, W.E. (2003) 'Advanced mixed methods research designs'. In Tashakkori, A. and Teddlie, C. (eds) *Handbook of Mixed Methods in Social and Behavioral Research*. Thousand Oaks, CA: SAGE Publications, 209–40.
- Denzin, N.K. (1970) *The Research Act: A theoretical introduction to sociological methods*. Chicago: Aldine.
- Gläser-Zikuda, M., Seidel, T., Rohlf, C., Gröschner, A. and Ziegelbauer, S. (eds) (2012) *Mixed Methods in der empirischen Bildungsforschung*. Münster: Waxmann.
- Jeismann, K.-E. (1997) 'Geschichtsbewusstsein – Theorie'. In Bergmann, K., Fröhlich, K., Kuhn, A., Rösen, J. and Schneider, G. (eds) *Handbuch der Geschichtsdidaktik*. 5th ed. Seelze: Kallmeyer, 42–5.
- Jeismann, K.-E. (2000) "'Geschichtsbewußtsein" als zentrale Kategorie der Didaktik des Geschichtsunterrichts'. In Jeismann, K.-E. *Geschichte und Bildung: Beiträge zur Geschichtsdidaktik und zur Historischen Bildungsforschung*. Ed. Jacobmeyer, W. and Schönemann, B. Paderborn: Schöningh, 46–72.
- Jeismann, K.-E., Kosthorst, E., Schäfer, B., Schlöder, B., Teppe, K. and Wasna, M. (1987) *Die Teilung Deutschlands als Problem des Geschichtsbewußtseins: Eine empirische Untersuchung über Wirkungen von Geschichtsunterricht auf historische Vorstellungen und politische Urteile*. Paderborn: Schöningh.
- Kelle, U. (2008) *Die Integration qualitativer und quantitativer Methoden in der empirischen Sozialforschung: Theoretische Grundlagen und methodologische Konzepte*. 2nd ed. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Kelle, U. (2014) 'Mixed methods'. In Baur, N. and Blasius, J. (eds), *Handbuch Methoden der empirischen Sozialforschung*. Wiesbaden: Springer, 153–66.
- Kintsch, W. and Van Dijk, T.A. (1978) 'Toward a model of text comprehension and production'. *Psychological Review*, 85 (5), 363–94.
- Kölbl, C. (2017) 'Geschichtsbewusstsein – Empirie'. In Barricelli, M. and Lücke, M. (eds) *Handbuch Praxis des Geschichtsunterrichts* (Vol. 1). 2nd ed. Schwalbach am Taunus: Wochenschau Verlag, 112–20.
- Körber, A. (2015) *Historical Consciousness, Historical Competencies – and Beyond? Some conceptual development within German history didactics*. Online. <https://tinyurl.com/yd5t5ttj> (accessed 28 December 2018).
- Körber, A., Schreiber, W. and Schöner, A. (eds) (2007) *Kompetenzen historischen Denkens: Ein Strukturmodell als Beitrag zur Kompetenzorientierung in der Geschichtsdidaktik*. Neuried: Ars Una.
- Köster, M. (2013) *Historisches Textverstehen: Rezeption und Identifikation in der multiethnischen Gesellschaft*. Berlin: Lit.
- Köster, M. (2016) 'Methoden empirischer Sozialforschung aus geschichtsdidaktischer Perspektive: Einleitung und Systematisierung'. In Thünemann, H. and Zülsdorf-Kersting, M. (eds) *Methoden geschichtsdidaktischer Unterrichtsforschung*. Schwalbach am Taunus: Wochenschau Verlag, 9–62.
- Kuckartz, U. (2014) *Mixed Methods: Methodologie, Forschungsdesigns und Analyseverfahren*. Wiesbaden: Springer.

- Lamnek, S. (2010) *Qualitative Sozialforschung: Lehrbuch*. 5th ed. Weinheim: Beltz.
- Mayring, P. (2001) 'Kombination und Integration qualitativer und quantitativer Analyse'. *Forum: Qualitative Sozialforschung*, 2 (1), Article 6, 1–14. Online. <http://nbn-resolving.de/urn:nbn:de:0114-fqs010162> (accessed 11 January 2018).
- Mayring, P. (2012) 'Mixed Methods – ein Plädoyer für gemeinsame Forschungsstandards qualitativer und quantitativer Methoden'. In Gläser-Zikuda, M., Seidel, T., Rohlf, C., Gröschner, A. and Ziegelbauer, S. (eds) *Mixed Methods in der empirischen Bildungsforschung*. Münster: Waxmann, 287–300.
- Merton, R.K. and Kendall, P.L. (1946) 'The focused interview'. *American Journal of Sociology*, 51 (6), 541–57.
- Morse, J.M. (1991) 'Approaches to qualitative-quantitative methodological triangulation'. *Nursing Research*, 40 (2), 120–3.
- Pandel, H.-J. (2005) *Geschichtsunterricht nach PISA: Kompetenzen, Bildungsstandards und Kerncurricula*. Schwalbach am Taunus: Wochenschau Verlag.
- Prinz, D. and Thünemann, H. (2016) 'Mixed-Methods-Ansätze in der empirischen Schul- und Unterrichtsforschung: Möglichkeiten und Grenzen für die Geschichtsdidaktik'. In Thünemann, H. and Zülsdorf-Kersting, M. (eds) *Methoden geschichtsdidaktischer Unterrichtsforschung*. Schwalbach am Taunus: Wochenschau Verlag, 229–53.
- Ross, A. (ed.) (1996) *Science Wars*. Durham, NC: Duke University Press.
- Rüsen, J. (1989) *Lebendige Geschichte: Grundzüge einer Historik III: Formen und Funktionen des historischen Wissens*. Göttingen: Vandenhoeck und Ruprecht.
- Schönemann, B. (2017) 'Geschichtsbewusstsein – Theorie'. In Barricelli, M. and Lücke, M. (eds) *Handbuch Praxis des Geschichtsunterrichts* (Vol. 1). 2nd ed. Schwalbach am Taunus: Wochenschau Verlag, 98–111.
- Seixas, P. (2016) 'Translation and its discontents: Key concepts in English and German history education'. *Journal of Curriculum Studies*, 48 (4), 427–39.
- Tashakkori, A. and Teddlie, C. (2003) 'The past and future of mixed methods research: From data triangulation to mixed model designs'. In Tashakkori, A. and Teddlie, C. (eds) *Handbook of Mixed Methods in Social and Behavioral Research*. Thousand Oaks, CA: SAGE Publications, 671–702.
- Teddlie, C. and Tashakkori, A. (2009) *Foundations of Mixed Methods Research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Thousand Oaks, CA: SAGE Publications.
- Thünemann, H. (2016) 'Probleme und Perspektiven der geschichtsdidaktischen Kompetenzdebatte'. In Handro, S. and Schönemann, B. (eds) *Aus der Geschichte lernen? Weiße Flecken der Kompetenzdebatte*. Berlin: Lit, 37–51.
- Trautwein, U., Bertram, C., Von Borries, B., Brauch, N., Hirsch, M., Klausmeier, K., Körber, A., Kühberger, C., Meyer-Hamme, J., Merkt, M., Neureiter, H., Schwan, S., Schreiber, W., Wagner, W., Waldis, M., Werner, M., Ziegler, B. and Zuckowski, A. (2017) *Kompetenzen historischen Denkens erfassen: Konzeption, Operationalisierung und Befunde des Projekts 'Historical Thinking – Competencies in History' (HiTCH)*. Münster: Waxmann.
- VanSledright, B.A. (2014) *Assessing Historical Thinking and Understanding: Innovative designs for new standards*. New York: Routledge.
- Von Borries, B., Fischer, C., Leutner-Ramme, S. and Meyer-Hamme, J. (2005) *Schulbuchverständnis, Richtlinienbenutzung und Reflexionsprozesse im Geschichtsunterricht: Eine qualitativ-quantitative Schüler- und Lehrerbefragung im deutschsprachigen Bildungswesen 2002*. Neuried: Ars Una.
- Waldis, M. and Ziegler, B. (eds) (2017) *Forschungswerkstatt Geschichtsdidaktik 15: Beiträge zur Tagung 'geschichtsdidaktik empirisch 15'*. Bern: hep.
- Werner, M. and Schreiber, W. (2015) 'Testfragen befragen – Pretesting und Optimierung des Large-Scale-Kompetenztests "HiTCH" durch Cognitive Labs'. In Waldis, M. and Ziegler, B. (eds) *Forschungswerkstatt Geschichtsdidaktik 13: Beiträge zur Tagung 'geschichtsdidaktik empirisch 13'*. Bern: hep, 153–64.
- Ziegler, B. and Hodel, J. (eds) (2009) *Forschungswerkstatt Geschichtsdidaktik 07: Beiträge zur Tagung 'geschichtsdidaktik empirisch 07'*. Bern: hep.