How to use mixed-methods and triangulation designs: An introduction to history education research

Udo Kelle – *University of the Federal Armed Forces Hamburg, Germany*
Christoph Kühberger – *University of Salzburg, Austria*
Roland Bernhard* – *University of Oxford, UK*

Abstract

As in many other social science disciplines, mixed methods and triangulation are gaining importance in history education research. Nevertheless, in this discipline there is also a prevailing lack of theoretical and methodological reflection about method integration. With this article, we wish to stimulate the methodological debate regarding this issue within the community of history education researchers and to strengthen the research profile of the discipline. We start by presenting lines of discussion regarding adequate research methods for the investigation of different types of social phenomena. Thereafter, we show how the 'paradigm wars' in social research were mitigated by the development of integrative concepts such as triangulation and mixed methods. Then we focus on current developments in history education research in German-speaking countries. Finally, we give a brief overview on international research into history teachers' beliefs, thereby addressing specific challenges for the application of triangulation or mixed methods in our discipline.

Keywords: mixed methods; triangulation; paradigm wars; teachers' beliefs; history education

Introduction

Empirical research on history education 'has grown dramatically in the past 35 years and exponentially in the last 15' (Epstein and Salinas, 2018: 61). History education is a growing area of enquiry conducted by researchers globally. Also, mixed-methods and triangulation designs are on the rise (to cite only a few mixed-methods studies in history education from the last three years: Bernhard and Kühberger, 2018; Rantala *et al.*, 2016; Baron, 2016; Harris and Burn, 2016; Cohen, 2016; Sant *et al.*, 2015; Yemini *et al.*, 2015). This movement towards mixed-methods or triangulation designs seems to reflect a general boom. It has been noted recently that for receiving competitive grants for research projects from funding agencies, a mixed-methods research design is not only very helpful but has nearly become a precondition, since there are 'expectations and requirements on the side of funding agencies that research in many fields should include a combination of quantitative and qualitative methods and a readiness to preferring such projects and proposals' (Flick, 2017: 1).

History education research deals with many elaborated and complex concepts such as 'epistemological beliefs', 'historical consciousness', 'historical thinking', 'historical culture', 'historical learning' and 'historical understanding'. The

discipline therefore includes many different aspects of research in various contexts. Mixed-methods research and triangulation are conceived as ways to account for such complexity (Ponce and Pagán-Maldonado, 2015) and to compensate for the methodological weaknesses of partial research approaches (McKim, 2017: 213; Gorard and Taylor, 2004: 4). We will argue in this introductory article that in our discipline there has not been much theoretical reflection about the application of triangulation and mixed methods so far. With this special edition, we thereby wish to stimulate the methodological discussion within the community and, in doing so, try to contribute to the sharpening and strengthening of the research profile of the discipline. Our paper also aims at overcoming emerging 'paradigm wars', which are perceived by some authors at this moment in history education research (see Köster and Thünemann in this special edition). Another aim of the Special Issue is to stimulate discussion between the German-speaking history education community and the international community.

In the first section of this introductory article, some lines of discussion about the question of which methods to use in research about social phenomena in general will be presented. In this context, strengths and weaknesses of the different paradigms will be examined. In the next section, we show how paradigm wars were overcome by promoting triangulation and mixed methods. Then, some lines of development within history education research in the German-speaking countries in the twentieth and twenty-first centuries are laid out. In the last section, we present a brief review of international research into history teachers' beliefs, also addressing the question of how triangulation and mixed-methods are dealt with in our discipline.

Paradigm wars: The methodological controversy in empirical social and educational research

Empirical research

The question of the 'correct method' for investigating the empirical world has sparked controversy in the social and educational sciences from the beginning of the twentieth century to the present day. This has primarily involved the question of the significance of empirical data and empirical research in general, and the relationship between empirical and non-empirical knowledge production. Educational researchers in particular have often emphasized the close links between their field and that of the classical humanities (see Kelle and Reith, 2014), and have kept a critical distance from an empiricist or 'positivist' understanding of science. On the other hand, successful advances have repeatedly been made in favour of decidedly empirically oriented educational research, such as the efforts made in the 1960s to achieve a 'realistic turn' (Roth, 1962) in Germany, or more recently the boom in empirical educational research following the large-scale comparative international studies TIMMS and PISA (see Lenzen *et al.*, 2004). Everyone entering the field of empirical research in the educational sciences and in history education is now inevitably drawn into a debate that has long occupied empirical social research in general: should we follow the path that scientific research has taken since the nineteenth century and attempt to grasp empirical reality primarily, or even exclusively, with the help of counting and measuring, quantifying, and experimental methods? Or does the subject matter of the social sciences require the application of specific non-standardized, open, qualitative methods?

Qualitative and quantitative research: Strengths and weaknesses

Social and educational science researchers are supposedly faced with a dichotomous decision: at first glance it seems as if one cannot do one thing without ignoring the other. The use of either qualitative or quantitative methods carries implications that seem to be hardly or not at all compatible with the postulates of the other tradition. This is the result of the fact that quantitative and qualitative methods have been developed in separate scientific communities to answer different research questions and to study different subjects. Divergent quality criteria and standards for good research have been developed in both traditions, which can easily come into conflict with each other.

Quantitative methods owe their importance to the interest in statistical distribution of certain characteristics and combinations of characteristics (including cognitive performance, specific competencies, personality traits, behavioural patterns, experiences of discrimination and victimization) in certain populations. Many of these phenomena only become visible when large groups are examined and compared with each other. A good example of this is the intergenerational transfer of educational inequality – the fact that children from socially and educationally disadvantaged homes have fewer opportunities than children from middle-class families to acquire advanced educational qualifications. This connection clearly exists, but it is by no means deterministic: some socially disadvantaged children also achieve a formally high or very high level of education, and, conversely, there are children from privileged backgrounds whose educational careers fail. When considering individual, or very few, cases, the likelihood of such cases coming to the fore is not small – in contrast, the exact opportunities for children from different social backgrounds to acquire educational qualifications can only be recorded statistically. An investigation of phenomena that require quantification is inconceivable without a standardization of the data collection process. From this, the well-known quality criteria of quantitative research (such as objectivity, reliability, validity and representativeness) are inevitably derived: when conducting such enquiries, investigators must ensure that all researchers can observe the same facts, that repeated measurements of the same facts produce the same results, that the measurement result actually says something about the facts examined and that the sample examined is sufficiently large and free of bias.

Standardization involves the methodical postulate of a hypothetico-deductive approach, which can turn out to be problematic in many research fields and for many research questions: to collect statistical information in the manner described, researchers must formulate hypotheses as precisely as possible and translate their theoretical concepts into measurement instruments (such as questionnaires or observation inventories) before collecting data. Quantitative research can only work properly if researchers develop theoretically founded and fairly precise ideas about what is going on in the field first, and subsequently use empirical data to test, empirically support or reject these ideas. But what happens when facts are investigated 'about which the researcher has no idea, because he has no in-depth knowledge of the relevant realm of reality'? These facts 'cannot … appear in his hypotheses at all, are therefore not tested either and [are] thus missing in the scientific image of the empirical field' (Gerdes, 1979: 5 [translated from German]. The fact that this critical question cannot simply be swept aside with the argument that a lack of theoretically derived hypotheses with empirical content is a sign of poor scientific practice has become clear in numerous works by cultural and social anthropologists, in which foreign societies or subcultures were investigated without quantitative methods and a strict hypothetico-deductive approach. The famous ethnographic studies carried out by Malinowski, Boas and

Mead, or the social science studies of the Chicago School in the urban ghettos of the 1920s would simply not have been feasible with questionnaire methods, as Whyte (2009) explains in the methodological appendix of his *Street Corner Society*. Especially, approaches of interpretative sociology (such as symbolic interactionism or phenomenological sociology) have provided theoretical and methodological justifications for social science researchers who collect empirical data not on the basis of *ex ante* formulated hypotheses but with the help of 'open' procedures, meaning without standardized instruments. Structures of meaning on the basis of which society members act and interact can be so fragmented and diverse in different (sub)cultures that researchers, who always have to rely on everyday knowledge in order to construct scientific instruments such as questionnaires (Cicourel, 1964; Kelle, 2008: 103), are not able to develop useful hypotheses and measurement instruments before an empirical investigation. Against this background, the reconstruction of structures of meaning with the help of qualitative data may become a much more important task of social research than the formulation and examination of hypotheses.

The paradigm wars

The models of research described here (hypothesis testing and quantification versus the reconstruction of meaning) from which different concepts and quality standards of 'good research' emerge seem to be incompatible at first glance. This has led to sometimes fierce controversies, and to the formation of independent, strongly differentiated scientific communities – a state that the educational researcher Gage (1989) has called 'paradigm wars'. In these wars, adherents of both traditions often reproached the other, sometimes with rather harsh words, on the basis of standards of good research developed in their own tradition.

The open (that is, non-standardized) collection of qualitative (unstructured or partly structured) data, imposes severe restrictions on the number of cases that can be dealt with for practical reasons alone. This has often provoked critical objections by quantitative social researchers: First, samples collected in qualitative research are often very small and usually not collected according to the usual methods and criteria of sampling developed in statistical sampling theory and survey research. Therefore, the generalizability of results in statistical terms often must remain dubious. Second, both the collection and the evaluation of qualitative data strongly depend on the person of the researcher. Such criticism was already formulated in the first textbooks of (quantitative) empirical social research. For example, in 1929 the statistician Lundberg criticized the fact that in many qualitative studies 'neither the validity of the sample nor of the interpretations are objectively demonstrable on account of the informality of the method' (Lundberg, 1942: 169). Despite decades of method development, which led to a differentiation and diversification of qualitative methods, this criticism remained at its core. For example, in a widely used German coursebook about quantitative methods in educational research, Wellenreuther (2000: 115) attacks a variety of qualitative studies, since they could not provide statistically representative results. Furthermore, the data (meaning interview excerpts) used in these studies would not be comparable, and would be evaluated based on a 'lord of the manor approach' (meaning that researchers selectively focused on material that supported their presuppositions) (ibid.: 316). According to Wellenreuther (ibid.: 14), there is 'often only sheer dilettantism' behind 'the mask of qualitative research', and serious qualitative research could at best only have a 'preparatory' function for quantitative empirical projects.

Criticisms of quantitative methods put forward by qualitative researchers are often similarly negative, or even pejorative and disdainful. Herbert Blumer (1969: 26),

one of the founding fathers of symbolic interactionism, came up with the verdict that quantitative methods such as quantitative surveys, statistical modelling or inferential statistical techniques are based on mere 'preoccupations and that much of present-day methodology in the social and psychological sciences is inadequate and misguided'. Another prominent proponent of the qualitative methodological tradition, William Filstead (1970: 3) stated that quantitatively oriented sociologists 'would bend, re-shape, and distort the empirical social world to fit the model they use to investigate it'.

Siegfried Lamnek's often quoted textbook of qualitative social research, which has become a classic in Germany, summarizes qualitative criticism of quantitative methods over more than 15 pages: qualitatively oriented social scientists accuse quantitative social research of representing a 'concept of a restricted experience' (Lamnek, 2005: 8 [translated from German]), producing 'ideological delusion' (ibid.: 9), being 'naive' and having a 'power-stabilizing' function (ibid.: 11), indulging in a 'measure fetishism' (ibid.: 12) and a 'false objectivity' (ibid.: 15), and imposing the 'researcher's perspective as a corset' (ibid.) on the actors in the research field.

Such polemics, which go as far as invectives, often carry a hardly concealed tendency to delegitimize opposing positions by stigmatizing them as unscientific and trying to exclude them from the realm of legitimate professional practice. As a consequence, qualitative and quantitative approaches have withdrawn themselves into segregated scientific communities, which have reduced possible areas of friction by refusing to discuss methodological issues with each other. Members of these communities publish in their own journals and manuals, organize their own conferences and meetings, and exchange information on the special problems of their approach in specific sections of academic societies.

Several authors try to make the continuing coexistence of two such controversial positions understandable by arguing that both approaches are based on different epistemological premises that are incompatible with each other. This idea draws on a thesis popularized by the philosopher and historian of science, Thomas Kuhn (2012), who claimed that scientific communities gather around their own 'paradigms', sets of basic assumptions that are protected by each group against criticism and modification. According to Kuhn, paradigms are incommensurable with each other because they are based on fundamentally different understandings of reality. Consequently, it would be fruitless to argue about their 'correctness' and appropriateness with rational and logically stringent arguments. The decision for or against a paradigm is, so to speak, made in a pre-rational space, and is comparable to a religious conversion.

With regard to qualitative and quantitative methods, these ideas were popularized by two American social scientists, Yvonna Lincoln and Egon Guba (Lincoln and Guba, 1985, 2000; Guba and Lincoln, 1994): quantitative and qualitative research styles are therefore based on certain basic epistemological tenets, which these authors called the 'positivist paradigm' and the 'constructivist paradigm', paradigms that are just as incompatible with each other as the views that the earth is a sphere or flat (Guba and Lincoln, 1988: 93). This idea can be criticized and attacked by drawing on a variety of epistemological and methodological arguments (see Kelle, 2008: 39; Tashakkori and Teddlie, 1998: 11; Morgan, 2007). However, the most striking argument against it is provided by research practice: since the beginnings of social research, empirical studies have been carried out time and again in which quantitative and qualitative methods are combined. These include several studies that have strongly promoted the methodological progress and theoretical development of the social sciences as a whole, which are cited very frequently and are still considered exemplary today, such as the 1929 'Middle Town study' by Robert and Helen Lynd (Lynd and Lynd, 1957),

the 'Marienthal study' on the consequences of unemployment by Marie Jahoda, Paul Lazarsfeld and colleagues (Jahoda *et al.*, 2017), the study on the 'authoritarian character' carried out in US exile by the Frankfurt Institute for Social Research (Adorno *et al.*, 1950) or the 'Hawthorne study' conducted by American industrial sociologists Fritz Roethlisberger and William Dickson (Roethlisberger *et al.*, 1939). But one does not need to look so far back into the past for a successful combination of qualitative and quantitative methods. Especially in the educational sciences, the number of studies that integrate qualitative and quantitative methods has multiplied considerably in recent decades, particularly stimulated by the new wave of empirical educational research (Rocco *et al.*, 2003; Gläser-Zikuda *et al.*, 2012; Mertens, 2014).

It is difficult to imagine that researchers in such studies experience a kind of intellectual schizophrenia in which they alternately (or simultaneously) believe in a flat or spherical earth (as the paradigm idea discussed above suggests), so the question arises as to how the obvious commensurability of quantitative and qualitative research in research practice can also be reflected and justified methodologically and epistemologically. There are now two prominent strands for discussion: one around the concept of methodological triangulation and the other around the idea of mixed-methods research. Both debates will be discussed briefly below.

Methodological triangulation

The term 'triangulation' was originally coined by two quantitative psychological methodologists: Donald Campbell and Donald Fiske (1959) used the term in their famous paper on 'multitrait–multimethod matrices' to offer an alternative to what they regarded as the 'simplified operationalism' of many psychologists. Both authors pleaded for the use of different measurement instruments to capture psychological traits and to determine the 'convergent' and 'discriminant validity' of the constructs thus operationalized by the extent of the statistical correlation between the results of the different measurement operations. This idea was later extended to the combination of different data classes and methods of data collection, in order to compensate for their respective validity restrictions (Webb *et al.*, 1966: 35). A qualitative social researcher, Norman Denzin, took up this idea in his monograph *The Research Act* in 1977, and expanded the concept by differentiating between different forms of triangulation, namely theory, data, methodological and investigator triangulation. The combination of qualitative and quantitative methods represented a possible type of methodological triangulation. Methodological triangulation would involve 'a complex process of playing each method off against the other in order to maximize the validity of field efforts' (Denzin, 1977: 310).

Several authors have criticized this idea by pointing out that research methods define their subject in different ways (others would say: they constitute it in the first place), and therefore may not be helpful to produce comparable results. Qualitative and quantitative methods often take their starting point from different theoretical and epistemological assumptions. If their results can then be related to each other, then it is rather in such a way that results become 'deeper' or 'broader' (which means better understandable or more comprehensive), without their validity necessarily being increased (see, above all, Fielding and Fielding, 1986: 33). Uwe Flick (1998: 230) puts it this way: 'Triangulation is less a strategy for validating results and procedures than an alternative to validation ... which increases scope, depth and consistency in methodological proceedings.'

One could therefore understand the concept of methodological triangulation in two different ways: as a validation of results by applying different methods, or as a combination of methods and/or data with the aim of describing a research field or topic more comprehensively and explaining it better with the help of different but complementary results (see also Erzberger and Kelle, 2003: 516; Kelle and Erzberger, 2004). If one examines examples of methodological triangulation from research practice, one will find that all three readings of the triangulation metaphor are applicable to describe specific instances of a combination of qualitative and quantitative methods:

1)　triangulation to examine the validity of (quantitative or qualitative) research results
2)　triangulation with the aim to better understand – through complementary findings – a result that could also exist on its own, or to place it in a broader context
3)　triangulation as a way of generating a complete result with the help of two partial findings that could not stand on their own.

The term 'triangulation' thus represents an ambiguous metaphor rather than a clear methodological concept. In fact, with regard to the combination of qualitative and quantitative methods, over the last two decades – especially in the Anglo-American language – it has increasingly been replaced by the term 'mixed methods'.

Mixed methods

Since the 2000s, the term 'mixed methods' has prevailed internationally over other terms for method integration, such as 'methodological triangulation', 'multi-methods', 'multi-methodologies', 'mixed research', 'mixed methodology' and 'method-plural research'. A movement began to emerge in the second half of the 1990s around the term 'mixed-method research', originally starting from educational and evaluation research in the USA, where paradigm wars were increasingly overcome by researchers who adopted pragmatic strategies to combine qualitative and quantitative methods in research practice. The publication of a monograph by Tashakkori and Teddlie (1998) on *Mixed Methodology* and a subsequent handbook on mixed methods in the social and behavioural sciences (Tashakkori and Teddlie 2003, 2010) marked the beginning of the formation of a vital scientific community. From its ranks, it has since founded both its own international academic society (Mixed Methods International Research Association) and established its own journal (*Journal of Mixed Methods Research*) with a good impact factor. Meanwhile the approach has gained international reputation and acceptance in the social and educational sciences, which can be also seen from the fact that sessions or streams on mixed methods have been held regularly for many years at major national and international conferences (such as those of the American Sociological Association, the International Sociological Association and the American Educational Research Association). In their work and debates, participants in the mixed-methods movement first strive for an appropriate epistemological and methodological justification for the integration of qualitative and quantitative research, and second, try to systematize the extensive field of strategies for combining qualitative and quantitative methods through appropriate taxonomies of 'mixed-method designs'.

An end to the paradigm wars: Mixed methods as a new research paradigm?

From the very beginning, representatives of the mixed-methods approach have countered the idea of a fundamental incommensurability and incompatibility of qualitative and quantitative methods, as claimed by adherents of the paradigm model,

with a pragmatic view: the appropriateness of research methods therefore cannot be decided solely on the basis of fundamental epistemological considerations, but only in relation to the concrete research questions and research field. Certain research questions and fields require the application of specific methods. Certain phenomena can only be empirically investigated if specific methods are used, as we have already shown with regard to the example of intergenerational transfer of educational inequality, which can hardly be made visible without statistical data. Conversely, many emergent cultural phenomena cannot be empirically investigated without the use of qualitative methods.

The question that follows from this, and that is currently under discussion, is whether this pragmatic orientation itself represents a 'third research paradigm' (Johnson and Onwuegbuzie, 2004: 14), based on corresponding philosophical doctrines as presented by Charles Peirce, John Dewey and William James in the context of American pragmatist philosophy (Morgan, 2007, Johnson *et al.*, 2017). This view has long been held by leading proponents of the mixed-methods movement, but has also been increasingly criticized (Hammersley, 2002; Maxwell, 2011; Kelle, 2017). These critics maintain that research methods are much less closely related to certain epistemological basic positions than is assumed by the paradigm model, which can be already seen from the fact that leading proponents of the same methodological traditions have often held divergent and even highly conflicting methodological and scientific theoretical positions (Maxwell and Loomis, 2003: 250). As alternatives to the idea that researchers must choose their methods primarily by drawing on a certain epistemological paradigm, it was proposed to regard epistemological ideas and concepts less as principles from which methodological research action can be directly derived, but rather as 'resources' (Seale, 1999: 26) or 'heuristics' (Maxwell, 2011), which may be used on a case-by-case basis and based on situational requirements. Such a pragmatic attitude towards research practice is certainly not the same as adherence to the philosophical school or 'paradigm' of American pragmatism. The pragmatic attitude rather follows the insight that successful social research is less a consequence of adopting abstract philosophical doctrines and principles than a consequence of mastery of craft techniques determined by practical and situational concerns (such as the expectations of colleagues and audiences). However, by adopting such a position, researchers may be seduced to neglect the importance of significant epistemological arguments (for example, about the relationship between correlation and causality) when evaluating research results. The shortcomings of the paradigm model can be overcome, nevertheless, if such epistemological arguments are not treated as the sole prerequisites and points of departure, but related to substantive theories and the concrete research questions and research fields (Kelle, 2017).

Mixed-methods designs

Even more than with epistemological questions, the mixed-methods approach has dealt with the systematization of strategies of method combination: in this area, a great body of terminology has now been developed, including a variety of different methodological concepts (each with its own history of terms and problems), not least of which is the term 'mixed methods' itself, which over time has prevailed over competing terms such as 'method triangulation'.

The starting point for this development was the task of identifying different functions of integration of qualitative and quantitative methods and then defining the designs that best fulfil these functions. This development started with a paper published by Greene and colleagues in 1989, which identified five major purposes of mixed-

methods research: triangulation (where this term only referred to mutual validation and corroboration, not to complementarity), 'complementarity', 'development', 'initiation' and 'expansion' (Greene *et al.*, 1989). Shortly after that, Morse proposed a notation system for mixed-methods designs, which became quite popular, and is now used (with different small amendments) in the majority of writings about mixed-methods research (Morse, 1991). However, the task of systematizing mixed-methods purposes and designs turned out to be much more challenging than initially thought. One may see that by realizing that the different functions of mixed methods suggested by Greene and her colleagues overlap: what is called 'development', for instance, can be used in some contexts as a means for triangulation; in others, it may be considered as a special type of complementarity. As a consequence, a great number of different typologies of mixed-methods designs have been proposed. In 2003, Creswell, Plano Clark, Gutmann and Hanson identified at least 15 different classifications in the literature (Creswell and Plano Clark, 2011). In 2003, Teddlie and Tashakkori noted that the question of the adequate typology of designs was a still unresolved issue in the mixed-methods debate at that time, and this is the case still today. The reason for that is that mixed-methods designs 'can be distinguished on a number of different dimensions … and developing a typology that would encompass all of those dimensions would be impossible' (Teddlie and Tashakkori, 2003: 26).

The most often used and seemingly most important dimensions refer to the chronological order in which the quantitative and qualitative parts of the design are implemented, on the one hand, and to the purpose of the study, on the other hand. With regard to chronological order, one can distinguish between sequential and parallel designs. In a sequential qualitative–quantitative design (if one employs the usual nomenclature, this would be called 'qual → quan'), a qualitative study is carried out, followed by a quantitative investigation; this order is reversed in quantitative–qualitative design (quan → qual). In a parallel design (quan + qual) quantitative and qualitative research is carried out simultaneously. This straightforward scheme gets complicated if one keeps in mind that in mixed-methods research, the qualitative and quantitative parts of a study often do not carry equal weight. Rather, one part can be more important for the respective research question and thus become dominant. This can be indicated by using small and big letters; for example, in a qual → QUAN design, first a small pilot study and thereafter a huge quantitative survey is carried out. Furthermore, the different designs can be (and often are) connected in different ways depending on the research question. This may result in a design in which one starts with a small qualitative pilot study and then conducts a first wave of a longitudinal quantitative panel study whereby also extended qualitative interviews are conducted with a subsample. Analysing the quantitative data, researchers decide to carry out even more qualitative interviews, before the second wave of the quantitative panel study takes place. The resulting design could be denoted as qual → QUAN + QUAL → QUAL → QUAN.

As far as the purposes of mixed-methods designs are concerned, it is only possible to provide an open-ended list of (partially overlapping) objectives of combining qualitative and quantitative methods. A mixed-methods design may fulfil one or more of the following functions: it may help to examine the scope and generalizability of qualitative findings through quantitative data; it may help to explain and understand otherwise incomprehensible quantitative results and unexplained variance by providing additional qualitative data; it may support the construction of quantitative measurement instruments or mitigate problems of quantitative measurement by helping to detect threats for validity and methodological artefacts;

it may assist qualitative sampling by providing information about the distribution of certain attributes in the investigated group; and it may help to close the gap between theory and empirical data in quantitative research by providing information about possible empirical references of abstract theoretical notions.

To simplify this complex picture, one may differentiate between three basic designs: (1) a sequential qual–quan design used to develop theoretical concepts, hypotheses or measurement instruments in a qualitative study, which are further elaborated and tested in a quantitative study; (2) a sequential quan–qual design that may assist with the explanation of difficult quantitative results, the detection and elucidation of quantitative measurement error and the development of qualitative sampling plans; and (3) a parallel qual–quan design, which can support the explanation of quantitative findings, the identification of threats for validity, and the generalization of qualitative findings.

Mixed methods and triangulation in history education research

To tie this general discussion back to history education research, it makes sense to trace the domain-specific developments. Based on research conducted in German-speaking countries, the aim is to make fundamental developments and tendencies visible. In order to gain a deeper insight into mixed methods and triangulation, important lines of development since 1950 will be illustrated, followed by a more in-depth example of international research on history teachers' beliefs.

General lines of development in the German-speaking countries

Intensive and heated discussions on methodological problems of empirical research are rare in history education in German-speaking countries. Only in the last decade has there been an increased interest and debate on various empirical approaches (see Hodel *et al.*, 2013; Waldis and Ziegler, 2017; Thünemann and Zülsdorf-Kersting, 2016; Prinz and Thünemann, 2016; Bernhard, 2018a; Bernhard and Kühberger, 2018; Bernhard, 2019). This is fundamentally related to an empirical turn within history education (Epstein and Salinas, 2018). Since the end of the twentieth century, various research projects have emerged in the wake of prominent educational research endeavours (PISA and PIRLS, for example). History education no longer primarily follows normative prescriptions or reduces its own field to pragmatics (teaching methodology), but rather has turned to differentiated empirical approaches that were used in the last century, enriched by current social science approaches. It is due to this increased emergence of empirical work, and to an internationally more interconnected scientific community (Ercikan and Seixas, 2015; Köster *et al.*, 2014), that even more complex research designs have found their way into the discussion and research on historical learning and thinking.

Looking at the development of history education in German-speaking countries, research settings in which triangulations have been implemented can be found, although they have not been labelled as such. One of the best-known and early examples is a study by Waltraud Küppers (1966). The German psychologist laid an important foundation for the study of historical learning in Germany with her studies on history teaching. Küppers investigated students from ages 4 to 12, and in her research, she conducted history lessons in 16 classes to gain a general impression of students' attitudes towards history. Afterwards, 40 classes were asked to write a total

of 1,400 essays on various topics. In order to triangulate these qualitative methods with quantitative approaches, Küppers designed a questionnaire of 50 knowledge items and three general questions on students' favourite subjects, their interest in history, and the source of their knowledge (Bracke *et al.*, 2014: 14–15). The author used methodological triangulation in order to supplement her qualitative surveys with quantitative data from a questionnaire. She does not label her approach as methodological triangulation, but nevertheless applies the concept itself: 'Eventually, however, it seemed necessary to emphasize the lines that suggested the lesson notes and essays. Since the methods mentioned allowed only qualitative processing, we sought to supplement these results with broader, quantitatively evaluated material (Küppers, 1966: 20 [translated from German]). Küppers notes a correlation between her qualitative observations in the classrooms, the essays and the results from the questionnaire, although she does not pinpoint, explain or reveal it in terms of the research design (ibid.: 20–1).

Another study cited more frequently today regarding its research design, is that of Jeismann *et al.* (1987). In this study, historians, history education researchers and psychologists attempted to combine data from open event and closed item questionnaires conducted with 653 students (ninth grade, age 14–15). Although the study, which was conceived as a complex multiphase design, carries some weaknesses (see Köster and Thünemann in this issue), the attempt to assess and correlate qualitative and quantitative data is worth mentioning.

The investigations by Küppers and Jeismann *et al.* are a good representation of research in German speaking countries in the twentieth century. At that time, the community of history education scholars was still leading a strong dialogue with psychological researchers. This tendency declined with the establishment of 'history didactics' as a new academic discipline during the second half of the twentieth century. Nevertheless, methodological reflection – especially from the perspective of history education – was not yet particularly pronounced and could only be tapped implicitly.

If one looks at studies from the year 2000 to the present, those of the German researcher Bodo von Borries and his team surely stand out (Von Borries *et al.*, 2005). In 2002, Bodo von Borries implemented a triangulation study on the use and comprehension of history textbooks. He and his team researched students (sixth, ninth and twelfth grades, ages 11–12, 14–15, 17–18), student teachers, and in-service teachers (NStud. = 1.291, NTeach. = 70) with a questionnaire. Items referred to knowledge, attitudes, emotions and competencies. Additionally, qualitative material (for example, essays and interviews) was gathered. Von Borries qualified his research setting as methodological triangulation, and addresses the complex interactions within his design between data from essays, post-interviews, quantitative questionnaires and so on in his paper (Von Borries, 2005a: 20–1) With regard to data synthesis, as it is presented in the research report by Von Borries *et al.*, it can be maintained that the researchers support a triangulation theory that tries to optimize the main results by using findings from different, but inter-related, aspects of the research setting. This kind of methodological triangulation brings together a variety of diverse qualitative and quantitative methodological approaches and their data in order to integrate different perspectives in one empirical explanation (Von Borries, 2005b: 269–304). The Swiss researcher Peter Gautschi, who works in the field of history education, presented a study on 'Good history lessons' in 2009. In his book, he includes several chapters on his own research and on the methodological discussion. Gautschi emphasizes that his research uses: (1) triangulation of data coming from classroom observation (videographies) and from questionnaires filled out by learners, teachers and experts; (2) a triangulation of different theoretical models applied simultaneously to validate

assumptions; and (3) a methodological triangulation that was reached by establishing a competition between qualitative and quantitative data in order to validate and enrich results (Gautschi, 2009: 124–5). According to the presentation of his results on 41 lessons, 689 students, 39 teachers and 10 experts, Gautschi recommended a triangulation design with laudatory words, since it would allow for the capturing of complex and multiple dependencies, especially in the context of history lessons (ibid.: 256).

Kühberger (2013, 2014) and his team researched Austrian students (n = 260) in the eighth grade (age 13), showing them a central scene of the movie *1492: Conquest of Paradise* (Ridley Scott/USA, 1994) about the landing of Columbus and his crew at Guanahani Island. They then asked the students to write an essay on whether the scene was authentically presented as it might have taken place in 1492. This quantitative survey was accompanied by two qualitative research methods in order to acquire complementary and more comprehensive results on the issue. The researchers thereby conducted additional interviews with some of the students after the writing of the essay on the same topic, and with teachers of the classes involved. This data triangulation, included explicitly in the study, provided a deeper insight into the phenomenon under scrutiny, and helped to investigate the different perspectives of the subjects involved (Kühberger, 2013, 2014). Quite a similar approach was implemented in a quantitative study on the handling of National Socialism, the Holocaust and memory culture in Austrian Schools (ninth grade, age 14). In this research, teachers and students received a quantitative questionnaire that was designed to reveal the views on teaching history via data triangulation from the two perspectives involved (that is, teachers and pupils) (Kühberger and Neureiter, 2017; Kühberger, 2017).

As the previous sections have shown, the concept of triangulation is still important for history education research at the beginning of the twenty-first century, especially as *terminus technicus*, whereas studies that explicitly include the terminology of mixed methods are still rare. Only very recently, a certain interest for the field of mixed methods can be noticed, as can be illustrated by this special issue. The Austrian project CAOHT, which inspired it, and in which the topics of triangulation and mixed methods were reflected upon (Bernhard, 2018a; Bernhard, 2019; Bernhard and Kühberger, 2018; Kühberger *et al.*, 2019; Bernhard *et al.*, 2019), will be described in more detail in the article by Bernhard in this issue.

Compared to the second half of the twentieth century, there is clearly more willingness today to explicitly discuss theoretical and practical questions of different methodological approaches. Recently, a new, albeit brief, attempt to systematize mixed-method approaches in the field of German-speaking history education was published (Prinz and Thünemann, 2016; see also Hasberg, 2004), and the notion of triangulation found its way into introductions to history education (Baumgärtner, 2015: 243). However, it would be presumptuous to claim that in all of the areas systematically presented in this paper, a sufficient domain-specific foundation of triangulation or mixed-methods design would be available that represents more than a transfer of approaches from the social or educational sciences. In the future, it will therefore be important to cultivate and systematize methodological reflections in history education, as well as create research projects that use triangulation or mixed-methods designs to obtain better research results on historical thinking and learning. Based on research activities in the area of history education, one can see an openness regarding rather different areas of triangulation and mixed-methods design.

Although an empirical turn in history education is clearly observable as an international trend, thus far there has been no 'dogma' in dealing with quantitative

or qualitative research methods, since both approaches are used side by side. This is certainly positive, but there is still the need for more discussion and reflection regarding different methodological approaches and their improvement. An international exchange about triangulation and mixed-methods designs would therefore be an important step forward that would help to systematize and further disseminate methodological knowledge already achieved in history education research.

Triangulation research in the field of history teachers' beliefs

A look at a specific research topic that is important for the history education community – namely, research about teachers' beliefs – would make it possible to gain exemplary insights into the ways in which triangulation and mixed methods are dealt with in the discipline. History teachers' beliefs are currently receiving rising international attention. In a careful inspection of more than forty English and German articles reporting such research from 1989 until 2017, we found 16 contributions in which a triangulation or mixed-methods approach was used. An analysis of the research designs in these articles reveals that 'within-' or 'between-method triangulation' designs are frequently used. Within-method triangulation designs, combining interviews and observations, were employed by Evans (1989, 1990), Wineburg and Wilson (1991), Schär and Sperisen (2011), and Thünemann (2012); Martell (2013) combined interviews, observations and teacher reflections; Virta (2002) combined interviews and essays; Barca *et al.* (2004) used essays and group discussions in their study; Wansink *et al.* (2017) used questionnaires with open-ended questions and performance tasks. Different studies with a between-methods triangulation design combined qualitative interviews and surveys. Evans (1989): quan → qual; Daumüller (2012): qual → quan; Maggioni (2010) used closed and open-ended questionnaires with teachers and students, 'constructed response tasks' and classroom observations. Fenn (2013) employed qualitative interviews and surveys to research how beliefs of history teachers can be changed (qual + quan). VanSledright and Reddy (2014) combined a quantitative questionnaire with qualitative interviews (quan → qual). Harris and Burn (2016) analysed both data from a quantitative questionnaire and written comments that the teachers made on this questionnaire (qual + quan) about the English history curriculum.

Is there a conscious reception of the discourse laid out in the articles about triangulation or mixed methods cited above, and what are the reasons given in these articles for using triangulation? To answer this question, all contributions cited above were analysed to answer the question whether the mixing of methods is somewhere explicitly mentioned or reflected upon. It turned out that, although the research designs of the analysed articles are often very elaborate and powerful, only 3 of the 16 analysed contributions made references to the debate about triangulation or mixed methods (Virta, 2002; Maggioni, 2010; Martell, 2013). Virta (2002) studied history student teachers' development and their beliefs about teaching and good teachers. She combined essays of 18 respondents and interviews with 5 student teachers; thus she used a within-method triangulation design, and she argues that in doing so she intended to increase the validity of her findings: 'One factor increasing validity of the data is the Triangulation carried out by complementing essay data with interviews that were conducted by another person' (Virta 2002: 697).

Maggioni (2010) used a variety of qualitative and quantitative methods to study epistemological beliefs of history teachers. She also argued that she intended to increase the validity of her study by using triangulation: 'In general, I have tried to

assess epistemic beliefs and historical thinking using a plurality of measures with the intent to triangulate results and address, as best as possible, issues of validity and reliability' (Maggioni, 2010: 116).

In a longitudinal study, Martell (2013) examined social studies teachers and their development of beliefs and practices related to teaching 'history as interpretation'. He described the combination of methods as follows:

> I also asked the beginning teachers in my study to write a reflection every other week during the first semester of their first year teaching. These written reflections allowed for Triangulation with interview and observation data by allowing me to acquire data more frequently, while allowing teachers to reflect on their practice. (Martell, 2013: 19)

These examples show that (although methods and methodologies are often mixed in history education), the international discourse about triangulation and mixing methods seems not to receive very much attention within this community. With this special edition of HERJ, we hope to inspire methodological awareness in this respect.

Summary

The preceding discussion should have made it clear that triangulation and mixed methods do not represent competing approaches or procedures. Rather, they are terms on different levels of meaning: triangulation generally refers to a combination of different elements (theories, methods, data or observer perspectives) in empirical research. Methodological triangulation as a special form of triangulation means the combination of different methods either with the aim of mutual validation of results, or with the goal to obtain a more adequate and comprehensive picture of the subject area through complementary results. Thereby, one may not only combine quantitative and qualitative methods, but also qualitative and quantitative methods among each other.

Mixed methods is a special form of methodological triangulation in which quantitative and qualitative methods are specifically combined. Since the 1990s, a very active scientific community has evolved with the aim to overcome the classical methodological divide between qualitative and quantitative research. In the past twenty years, there has been decisive progress in the epistemological and methodological debate in the social and educational sciences, as well as with regard to the development of research in social and educational research methods.

History education has undergone its own development within these general strands. In this field, an intensifying debate and growing awareness of the possibilities of mixed methods and triangulation can be observed that hopefully may be further stimulated by this special edition. The traditional polar positions in the paradigm wars – claiming the fundamental superiority of one´s own approach and the lack of scientific rigour or seriousness on the other side – turned out to lack sound philosophic ground and failed to demonstrate practical sustainability in the long run. Consequently, there was space for the idea that there is not one right approach to empirical research. Rather, both traditions show specific strengths and weaknesses. Fortunately, these are often complementary, so that the weaknesses of one tradition can be identified, worked on and balanced by drawing on the strengths of the other. Both qualitative and quantitative methods have to struggle with threats to validity, limits of knowledge and methodological problems, which nevertheless can be addressed with the help of the other tradition. This approach now invites us to give an honest account of the limitations of our research methods and methodologies, and to do research for their improvement

and further development. A lot of work has still to be done to systematically record and describe problems of qualitative and quantitative research in specific subject areas in order to develop concrete mixed-method designs on this basis. This will certainly demand more and more close cooperation between methodologists and empirical researchers in different social and educational science disciplines. We very much hope that this volume will be helpful in enabling such a productive dialogue.

Notes on the contributors

Udo Kelle is Professor for Social Research Methods and Statistics at the Helmut Schmidt University of the Federal Armed Forces in Hamburg. He has taught social research methods at universities in Germany (including Marburg and Bremen) and abroad (including Vienna and Oslo). He has published several books and more than a hundred articles about qualitative and quantitative methods and their integration in social research.

Christoph Kühberger is Professor for History and Civic Education at the University of Salzburg (Austria). He was Professor for European Cultural History at the University of Hildesheim (Germany) and Professor for History and Civic Education at the Pedagogical University of Salzburg Stefan Zweig (Austria). His current research interests are history education and civic education, ethnography, historical culture, new cultural history and the ethics of historical sciences.

Roland Bernhard is a visiting research fellow at the Department of Education at the University of Oxford (UK), and a post-doctoral researcher and lecturer in history and civic education at the University of Salzburg (Austria). He has taught history education and cultural studies at several universities in Austria (Graz, Innsbruck, Salzburg), and has published a range of books and articles, mainly in the area of history education. He worked for several years as a teacher in schools.

References

Adorno, T.W., Frenkel-Brunswik, E., Levinson, D.J. and Sanford, R.N. (1950) *The Authoritarian Personality*. Online. www.ajcarchives.org/main.php?GroupingId=6490 (accessed 23 January 2017).

Barca, I., Magalhães, O. and Castro, J. (2004) 'Ideas on history and orientation in time: A study with beginning teachers'. *International Journal of Historical Learning, Teaching and Research*, 4 (2), 37–43.

Baron, C. (2016) 'Using embedded visual coding to support contextualization of historical texts'. *American Educational Research Journal*, 53 (3), 516–40.

Baumgärtner, U. (2015) *Wegweiser Geschichtsdidaktik: Historisches Lernen in der Schule*. Paderborn: Schöningh.

Bernhard, R. (2018a) 'Fragebogenentwicklung anhand qualitativer Daten in einem Mixed-Methods-Research-Design: Eine geschichtsdidaktische Perspektive zu historischem Denken und Schulbuchnutzung'. In Bramann, C., Kühberger, C. and Bernhard, R. (eds) *Historisch Denken lernen mit Schulbüchern*. Schwalbach am Taunus: Wochenschau Verlag, 37–62.

Bernhard, R. (2018b) 'Teaching to think historically using textbooks: Insights for initial teacher education drawn from a qualitative empirical study in Austria'. *Educatio Siglo XXI*, 36 (1), S39–56.

Bernhard, R. (2019) 'Das Schulbuch als Leitmedium des Geschichtsunterrichts in Österreich. Empirische Ergebnisse einer Triangulationsstudie und einige Schlussfolgerungen für die LehrerInnenbildung'. In Bernhard, R., Bramann, C. and Kühberger, C. (eds.) *Das Geschichtsschulbuch: Lernen – Lehren – Forschen*. Münster/New York: Waxmann (forthcoming).

Bernhard, R., Bramann, C. and Kühberger, C. (2019) *Verwendung des Geschichtsschulbuches durch Schüler/innen und Lehrer/innen: Empirische Hauptergebnisse des Mixed-Method-Projektes CAOHT* (in preparation).

Bernhard, R. and Kühberger, C. (2018) 'Erforschung der Einstellungen von Geschichtslehrpersonen zu fachspezifischer Kompetenzorientierung. Einige Ergebnisse und theoretische Reflexionen zu Mixed Methods'. In Schreiber, W., Ziegler, B. and Kühberger, C. (eds.) *Geschichtsdidaktischer Zwischenhalt. Beiträge aus der Tagung Kompetent machen für ein Leben in, mit und durch Geschichte in Eichstätt vom November 2017*. Münster: Waxmann, 145–58.

Blumer, H. (1969) *Symbolic Interactionism: Perspective and method*. Englewood Cliffs, NJ: Prentice-Hall.

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.

Campbell, D.T. and Fiske, D.W. (1959) 'Convergent and discriminant validation by the multitrait-multimethod matrix'. *Psychological Bulletin*, 56 (2), 81–105.

Cicourel, A.V. (1964) *Method and Measurement in Sociology*. New York: Free Press of Glencoe.

Cohen, E.H. (2016) 'Teacher autonomy within a flexible national curriculum: Development of Shoah (Holocaust) education in Israeli state schools'. *Journal of Curriculum Studies*, 48 (2), 167–91.

Creswell, J.W. and Plano Clark, V.L. (2011) *Designing and Conducting Mixed Methods Research*. 2nd ed. Los Angeles: SAGE Publications.

Daumüller, M. (2012) 'Einstellungen und Haltungen von Fachlehrerinnen und Fachlehrern'. In Barricelli, M. and Lücke, M. (eds) *Handbuch Praxis des Geschichtsunterrichts* (Vol. 2). Schwalbach am Taunus: Wochenschau Verlag, 370–85.

Denzin, N.K. (1977) *The Research Act: A theoretical introduction to sociological methods*. New York: McGraw-Hill.

Epstein, T. and Salinas, C.S. (2018) 'Research methodologies in history education'. In Metzger, S.A. and McArthur Harris, L. (eds) *The Wiley International Handbook of History Teaching and Learning*. Hoboken, NJ: Wiley-Blackwell, 61–91.

Ercikan, K. and Seixas, P. (eds) (2015) *New Directions in Assessing Historical Thinking*. New York: Routledge.

Erzberger, C. and Kelle, U. (2003) 'Making inferences in mixed methods: The rules of integration'. In Tashakkori, A. and Teddlie, C. (eds) *Handbook of Mixed Methods in Social and Behavioral Research*. Thousand Oaks, CA: SAGE Publications, 457–90.

Evans, R.W. (1989) 'Teacher conceptions of history'. *Theory and Research in Social Education*, 17 (3), 210–40.

Evans, R.W. (1990) 'Teacher conceptions of history revisited: Ideology, curriculum, and student belief'. *Theory and Research in Social Education*, 18 (2), 101–38.

Fenn, M. (2013) 'Konstruktivistisches Geschichtsverständnis im Unterricht fördern: Studierende ändern ihr Lehrverhalten von einseitig instruktional in problemorientiert'. In Hodel, J., Waldis, M. and Ziegler, B. (eds) *Forschungswerkstatt Geschichtsdidaktik 12: Beiträge zur Tagung "geschichtsdidaktik empirisch 12"*. Bern: hep, 60–71.

Fielding, N.G. and Fielding, J.L. (1986) *Linking Data* (Qualitative Research Methods 4). Beverly Hills, CA: SAGE Publications.

Filstead, W.J. (1970) 'Introduction'. In Filstead, W.J. (ed.) *Qualitative Methodology: Firsthand involvement with the social world*. Chicago: Markham Publishing Company, 1–11.

Flick, U. (1998) *An Introduction to Qualitative Research*. London: SAGE Publications.

Flick, U. (2017) 'Mantras and myths: The disenchantment of mixed-methods research and revisiting triangulation as a perspective'. *Qualitative Inquiry*, 23 (1), 46–57.

Gage, N.L. (1989) 'The paradigm wars and their aftermath: A "historical" sketch of research on teaching since 1989'. *Educational Researcher*, 18 (7), 4–10.

Gautschi, P. (2009) *Guter Geschichtsunterricht: Grundlagen, Erkenntnisse, Hinweise*. Schwalbach am Taunus: Wochenschau Verlag.

Gerdes, K. (ed.) (1979) *Explorative Sozialforschung*. Stuttgart: Enke.

Gläser-Zikuda, M., Seidel, T., Rohlfs, C., Gröschner, A. and Ziegelbauer, S. (eds) (2012) *Mixed Methods in der Empirischen Bildungsforschung*. Münster: Waxmann.

Gorard, S. and Taylor, C. (2004) *Combining Methods in Educational and Social Research*. Maidenhead: Open University Press.

Greene, J.C., Caracelli, V.J. and Graham, W.F. (1989) 'Toward a conceptual framework for mixed-method evaluation designs'. *Educational Evaluation and Policy Analysis*, 11 (3), 255–74.

Guba, E.G. and Lincoln, Y.S. (1988) 'Do inquiry paradigms imply inquiry methodologies?' In Fetterman, D.M. (ed.) *Qualitative Approaches to Evaluation in Education: The silent scientific revolution*. New York: Praeger, 88–115.

Guba, E.G. and Lincoln, Y.S. (1994) 'Competing paradigms in qualitative research'. In Denzin, N.K. and Lincoln, Y.S. (eds) *Handbook of Qualitative Research*. Thousand Oaks, CA: SAGE Publications, 105–17.

Hammersley, M. (2002) 'The relationship between qualitative and quantitative research: Paradigm loyalty versus methodological eclecticism'. In Richardson, J.T.E. (ed.) *Handbook of Qualitative Research Methods for Psychology and the Social Sciences*. Oxford: BPS Blackwell, 159–74.

Harris, R. and Burn, K. (2016) 'English history teachers' views on what substantive content young people should be taught'. *Journal of Curriculum Studies*, 48 (4), 518–46.

Hasberg, W. (2004) 'Kulturvergleichende Richtlinienanalyse: Triangulation als Notwendigkeit und Weg'. In Handro, S. and Schönemann, B. (eds) *Geschichtsdidaktische Lehrplanforschung: Methoden – Analysen – Perspektiven*. Münster: Lit, 27–50.

Hodel, J., Waldis, M. and Ziegler, B. (eds) (2013) *Forschungswerkstatt Geschichtsdidaktik 12: Beiträge zur Tagung "geschichtsdidaktik empirisch 12"*. Bern: hep.

Jahoda, M., Lazarsfeld, P.F. and Zeisel, H. (2017) *Marienthal: The sociography of an unemployed community*. New York: Routledge.

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.

Johnson, R.B., de Waal, C., Stefurak, T. and Hildebrand, D.L. (2017) 'Understanding the philosophical positions of classical and neopragmatists for mixed methods research'. In Baur, N., Kelle, U. and Kuckartz, U. (eds) *Mixed Methods* (Kölner Zeitschrift für Soziologie und Sozialpsychologie Sonderheft 57). Wiesbaden: Springer VS, 63–86.

Johnson, R.B. and Onwuegbuzie, A.J. (2004) 'Mixed methods research: A research paradigm whose time has come'. *Educational Researcher*, 33 (7), 14–26.

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. (2017) 'Die Integration qualitativer und quantitativer Forschung – theoretische Grundlagen von "Mixed Methods"'. In Baur, N., Kelle, U. and Kuckartz, U. (eds) *Mixed Methods* (Kölner Zeitschrift für Soziologie und Sozialpsychologie Sonderheft 57). Wiesbaden: Springer VS, 39–61.

Kelle, U. and Erzberger, C. (2004) 'Qualitative and quantitative methods: Not in opposition'. In Flick, U., Von Kardorff, E. and Steinke, I. (eds) *A Companion to Qualitative Research*. London: SAGE Publications, 172–7.

Kelle, U. and Reith, F. (2014) 'Empirisch forschen in der Pädagogik – qualitativ und/oder quantitativ?'. In Ziegler, A. and Zwick, E. (eds) *Theoretische Perspektiven der modernen Pädagogik*. Münster: Lit, 25–44.

Köster, M., Thünemann, H. and Zülsdorf-Kersting, M. (eds) (2014) *Researching History Education: International perspectives and disciplinary traditions*. Schwalbach am Taunus: Wochenschau Verlag.

Kühberger, C. (ed.) (2013) *Geschichte denken: Zum Umgang mit Geschichte und Vergangenheit von Schüler/innen der Sekundarstufe I am Beispiel "Spielfilm": Empirische Befunde – Diagnostische Tools – Methodische Hinweise*. Innsbruck: StudienVerlag.

Kühberger, C. (2014) 'Empirische Befunde zum Umgang mit Spielfilmen über die Vergangenheit in der Sekundarstufe I'. *Geschichte in Wissenschaft und Unterricht*, 7–8, 423–38.

Kühberger, C. (2017) 'Teaching the Holocaust and national socialism in Austria: Politics of memory, history classes, and empirical insights'. *Holocaust Studies: A Journal of Culture and History*, 23 (3), 396–424.

Kühberger, C., Bernhard, R. and Bramann, C. (eds) (2019) *Das Geschichtsschulbuch: Lernen – Lehren – Forschen*. Münster/New York: Waxmann.

Kühberger, C. and Neureiter, H. (2017) *Zum Umgang mit Nationalsozialismus, Holocaust und Erinnerungskultur: Eine quantitative Untersuchung bei Lernenden und Lehrenden an Salzburger Schulen aus geschichtsdidaktischer Perspektive*. Schwalbach am Taunus: Wochenschau Verlag.

Kuhn, T.S. (2012) *The Structure of Scientific Revolutions*. 4th ed. Chicago: University of Chicago Press.

Küppers, W. (1966) *Zur Psychologie des Geschichtsunterrichts*. 2nd ed. Stuttgart: Klett.

Lamnek, S. (2005) *Qualitative Sozialforschung: Lehrbuch*. 4th ed. Weinheim: Beltz.

Lenzen, D., Baumert, J., Watermann, R. and Trautwein, U. (eds) (2004) *PISA und die Konsequenzen für die erziehungswissenschaftliche Forschung*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Lincoln, Y.S. and Guba E.G. (1985) *Naturalistic Inquiry*. Newbury Park, CA: SAGE Publications.

Lincoln, Y.S. and Guba E.G. (2000) 'Paradigmatic controversies, contradictions, and emerging confluences'. In Denzin, N.K. and Lincoln, Y.S. (eds) *Handbook of Qualitative Research*. 2nd ed. Thousand Oaks, CA: SAGE Publications, 163–88.

Lundberg, G.A. [1929] (1942) *Social Research: A study in methods of gathering data*. New York: Longmans, Green and Company.

Lynd, R.S. and Lynd, H.M. [1929] (1957) *Middletown: A study in American culture*. San Diego: Harcourt.

Maggioni, L. (2010) 'Studying Epistemic Cognition in the History Classroom: Cases of teaching and learning to think historically'. Unpublished PhD thesis, University of Maryland.

Martell, C.C. (2013) 'Learning to teach history as interpretation: A longitudinal study of beginning teachers'. *Journal of Social Studies Research*, 37 (1), 17–31.

Maxwell, J.A. (2011) 'Paradigms or toolkits? Philosophical and methodological positions as heuristics for mixed methods research'. *Mid-Western Educational Researcher*, 24 (2), 27–30.

Maxwell, J.A. and Loomis, D.M. (2003) 'Mixed methods design: An alternative approach'. In Tashakkori, A. and Teddlie, C. (eds) *Handbook of Mixed Methods in Social and Behavioral Research*. Thousand Oaks, CA: SAGE Publications, 241–72.

McKim, C.A. (2017) 'The value of mixed methods research: A mixed methods study'. *Journal of Mixed Methods Research*, 11 (2), 202–22.

Mertens, D.M. (2014) *Research and Evaluation in Education and Psychology: Integrating diversity with quantitative, qualitative, and mixed methods*. 4th ed. Thousand Oaks, CA: SAGE Publications.

Morgan, D.L. (2007) 'Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods'. *Journal of Mixed Methods Research*, 1 (1), 48–76.

Morse, J.M. (1991) 'Approaches to qualitative-quantitative methodological triangulation'. *Nursing Research*, 40 (2), 120–3.

Ponce, O.A. and Pagán-Maldonado, N. (2015) 'Mixed methods research in education: Capturing the complexity of the profession'. *International Journal of Educational Excellence*, 1 (1), 111–35.

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.

Rantala, J., Manninen, M. and Van den Berg, M. (2016) 'Stepping into other people's shoes proves to be a difficult task for high school students: Assessing historical empathy through simulation exercise'. *Journal of Curriculum Studies*, 48 (3), 323–45.

Rocco, T.S., Bliss, L.A., Gallagher, S., Perez-Prado, A., Alacaci, C., Dwyer, E.S., Fine, J.C. and Pappamihiel, N.E. (2003) 'The pragmatic and dialectical lenses: Two views of mixed methods use in education'. In Tashakkori, A. and Teddlie, C. (eds) *Handbook of Mixed Methods in Social and Behavioral Research*. Thousand Oaks, CA: SAGE Publications, 595–615.

Roethlisberger, F.J., Dickson, W.J. and Wright, H.A. (1939) *Management and the Worker*. Cambridge, MA: Harvard University Press.

Roth, H. (1962) 'Die realistische Wendung in der Pädagogischen Forschung'. *Neue Sammlung*, 2, 482–90.

Sant, E., González-Monfort, N., Santisteban Fernández, A., Pagès Blanch, J. and Oller Freixa, M. (2015) 'How do Catalan students narrate the history of Catalonia when they finish primary education?'. *McGill Journal of Education*, 50 (2–3), 341–62.

Schär, B.C. and Sperisen, V. (2011) 'Zum Eigensinn von Lehrpersonen im Umgang mit Lehrbüchern: Das Beispiel "Hinschauen und Nachfragen"'. In Hodel, J. and Ziegler, B. (eds) *Forschungswerkstatt Geschichtsdidaktik 09: Beiträge zur Tagung "geschichtsdidaktik empirisch 09"*. Bern: hep, 124–34.

Seale, C. (1999) *The Quality of Qualitative Research*. London: SAGE Publications.

Tashakkori, A. and Teddlie, C. (1998) *Mixed Methodology: Combining qualitative and quantitative approaches*. Thousand Oaks, CA: SAGE Publications.

Tashakkori, A. and Teddlie, C. (eds) (2003) *Handbook of Mixed Methods in Social and Behavioral Research*. Thousand Oaks, CA: SAGE Publications.

Tashakkori, A. and Teddlie, C. (eds) (2010) *SAGE Handbook of Mixed Methods in Social and Behavioral Research*. 2nd ed. Thousand Oaks, CA: SAGE Publications.

Teddlie, C. and Tashakkori, A. (2003) 'Major issues and controversies in the use of mixed methods in the social and behavioral sciences'. In Tashakkori, A. and Teddlie, C. (eds) *Handbook of Mixed Methods in Social and Behavioral Research*. Thousand Oaks, CA: SAGE Publications, 3–50.

Thünemann, H. (2012) '"Ganz, ganz historisch gedacht": Merkmale guten Geschichtsunterrichts aus Lehrerperspektive'. In Meyer-Hamme, J., Thünemann, H. and Zülsdorf-Kersting, M. (eds) *Was heißt guter Geschichtsunterricht? Perspektiven im Vergleich*. Schwalbach am Taunus: Wochenschau Verlag, 39–54.

Thünemann, H. and Zülsdorf-Kersting, M. (eds) (2016) *Methoden geschichtsdidaktischer Unterrichtsforschung*. Schwalbach am Taunus: Wochenschau Verlag.

VanSledright, B. and Reddy, K. (2014) 'Changing epistemic beliefs? An exploratory study of cognition among prospective history teachers'. *Revista Tempo e Argumento*, 6 (11), 28–68.

Virta, A. (2002) 'Becoming a history teacher: Observations on the beliefs and growth of student teachers'. *Teaching and Teacher Education*, 18 (6), 687–98.

Von Borries, B. (2005a) 'Fragestellung und Verfahrensweise'. In Von Borries, B., Fischer, C., Leutner-Ramme, S. and Meyer-Hamme, J. *Schulbuchverständnis, Richtlinienbenutzung und Reflexionsprozesse im Geschichtsunterricht: Eine qualitativ-quantitative Schüler- und Lehrerbefragung im Deutschsprachigen Bildungswesen 2002*. Bayerische Studien zur Geschichtsdidaktik, Vol. 9. Neuried: Ars Una, 11–36.

Von Borries, B. (2005b) 'Synthese und Folgerungen'. In Von Borries, B., Fischer, C., Leutner-Ramme, S. and Meyer-Hamme, J. *Schulbuchverständnis, Richtlinienbenutzung und Reflexionsprozesse im Geschichtsunterricht: Eine qualitativ-quantitative Schüler- und Lehrerbefragung im Deutschsprachigen Bildungswesen 2002*. Bayerische Studien zur Geschichtsdidaktik, Vol. 9. Neuried: Ars Una, 269–304.

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*. Bayerische Studien zur Geschichtsdidaktik, Vol. 9. Neuried: Ars Una.

Waldis, M. and Ziegler, B. (eds) (2017) *Forschungswerkstatt Geschichtsdidaktik 15: Beiträge zur Tagung "geschichtsdidaktik empirisch 15"*. Bern: hep.

Wansink, B.G.J., Akkerman, S.F., Vermunt, J.D., Haenen, J.P.P. and Wubbels, T. (2017) 'Epistemological tensions in prospective Dutch history teachers' beliefs about the objectives of secondary education'. *Journal of Social Studies Research*, 41 (1), 11–24.

Webb, E.J., Campbell, D.T., Schwartz, R.D. and Sechrest, L. (1966) *Unobtrusive Measures: Nonreactive research in the social sciences*. Chicago: Rand McNally.

Wellenreuther, M. (2000) *Quantitative Forschungsmethoden in der Erziehungswissenschaft: Eine Einführung*. Weinheim: Juventa.

Whyte, W.F. (2009) *Street Corner Society: The social structure of an Italian slum*. 4th ed. Chicago: University of Chicago Press.

Wineburg, S.S. and Wilson, S.M. (1991) 'Models of wisdom in the teaching of history'. *The History Teacher*, 24 (4), 395–412.

Yemini, M., Yardeni-Kuperberg, O. and Natur, N. (2015) 'The global–local nexus: Desired history curriculum components from the perspective of future teachers in a conflict-ridden society'. *Journal of Curriculum Studies*, 47 (1), 26–48.