Teachers trust educational science - Especially if it confirms their beliefs

Kirstin Schmidt[1]*, Tom Rosman[2], Colin Cramer[3],
Kris-Stephen Besa[4] and Samuel Merk[1]

[1]Faculty of Humanities and Human Science, Institute for School and Classroom Development, Karlsruhe University of Education, Karlsruhe, Germany, [2]Department for Research Literacy and User-Friendly Research Support, Leibniz Institute for Psychology, Trier, Germany, [3]Faculty of Economics and Social Sciences, Institute of Education, University of Tübingen, Tübingen, Germany, [4]Institute for Educational Science, Educational Science and Social Science, University of Münster, Münster, Germany

Teachers around the world are increasingly required by policy guidelines to inform their teaching practices with scientific evidence. However, due to the division of cognitive labor, teachers often cannot evaluate the veracity of such evidence first-hand, since they lack specific methodological skills, such as the ability to evaluate study designs. For this reason, second-hand evaluations come into play, during which individuals assess the credibility and trustworthiness of the person or other entity who conveys the evidence instead of evaluating the information itself. In doing so, teachers' belief systems (e.g., beliefs about the trustworthiness of different sources, about science in general, or about specific educational topics) can play a pivotal role. But judging evidence based on beliefs may also lead to distortions which, in turn, can result in barriers for evidence-informed school practice. One popular example is the so-called confirmation bias, that is, preferring belief-consistent and avoiding or questioning belief-inconsistent information. Therefore, we experimentally investigated (1) whether teachers trust knowledge claims made by other teachers and scientific studies differently, (2) whether there is an interplay between teachers' trust in these specific knowledge claims, their trust in educational science, and their global trust in science, and (3) whether their prior topic-specific beliefs influence trust ratings in the sense of a confirmation bias. In an incomplete rotated design with three preregistered hypotheses, $N = 414$ randomly and representative sampled in-service teachers from Germany indicated greater trust in scientific evidence (information provided by a scientific journal) compared to anecdotal evidence (information provided by another teacher on a teacher blog). In addition, we found a positive relationship between trust in educational science and trust in specific knowledge claims from educational science. Finally, participants also showed a substantial confirmation bias, as they trusted educational science claims more when these matched (rather than contradicted) their prior beliefs. Based on these results, the interplay of trust, first-hand evaluation, and evidence-informed school practice is discussed.

KEYWORDS

evidence-informed education, anecdotal evidence, confirmation bias, second-hand evaluation, scientific evidence, trust in science, teacher education

1. Introduction

Teachers can inform their professional practice using a vast number of information sources. To name just a few, they can refer to their own teaching experience, they can follow the advice of their colleagues, or they may refer to evidence obtained through educational science (Buehl and Fives, 2009). Around the world, policymakers (e.g., European Commission, 2007; Kultusministerkonferenz, 2014) as well as scientists (e.g., Bauer and Prenzel, 2012) increasingly value the latter and often consider it the most veracious body of knowledge because it is obtained through systematic and verifiable means (e.g., Williams and Coles, 2007; Bauer and Prenzel, 2012; Bauer et al., 2015; Brown et al., 2017). Such so-called evidence-informed school practice is considered to have large potential to improve school and teaching quality as well as student learning, for example by helping teachers to (a) make or change decisions pertaining to their teaching practice, (b) develop new practices, (c) inform leadership (Brown, 2020; Brown et al., 2022), or (d) effectively deal with problems that repeatedly come up in practice (Stark, 2017; Kiemer and Kollar, 2021). Empirical studies provide initial evidence for these considerations by showing that scientific evidence can actually inform teachers' practical decisions and actions (Cain, 2015), and that evidence-informed interventions, at least in specific contexts like formative assessment (Black and Wiliam, 2003) or mathematics learning (Doabler et al., 2014), can positively affect student achievement.

However, transforming scientific evidence into educational practice can be a challenging endeavor. Educational stakeholders increasingly take the position that scientific evidence can only *enrich* teachers' practical experience and contextual knowledge (Brown et al., 2017; Kiemer and Kollar, 2021). Thus, scientific evidence may not provide teachers with action-guiding recipes that can be directly used in everyday practice. Rather, it can be used to stimulate, reflect, or revise professional actions and decisions, which additionally makes them more transparent and objective (e.g., Bauer et al., 2015; Brown et al., 2017). In addition, considering evidence enables teachers to rationally justify their decisions and actions and consequently make them explicit—even to others, such as colleagues or parents (Bauer et al., 2015). The complexity of this endeavor is also reflected in the implementation steps of evidence-informed school practice. Evidence-informed practices require educational research literacy (e.g., Shank and Brown, 2007) which not only includes (1) accessing, (2) comprehending, and (3) critically reflecting the scientific evidence (e.g., its validity), but also (4) combining this evidence with prior knowledge before (5) using the evidence in practice (Shank and Brown, 2007; Brown et al., 2022).

This sophisticated theoretical view on teachers' professional use of scientific evidence contradicts reality: Even though policy guidelines all over the world (e.g., Kultusministerkonferenz, 2014 in Germany) emphasize the importance of evidence-informed school practice, it becomes apparent that teachers often do not follow such guidelines (e.g., Brown et al., 2017; Hinzke et al., 2020). Hence, the question arises what barriers teachers might face when it comes to realizing evidence-informed practice.

Previous studies identified various barriers ranging from lack of time to engage with evidence on top of other professional tasks, belief systems that devalue the importance, practicality, and usability of scientific evidence for teaching practice, as well as a lack of skills or knowledge to evaluate scientific evidence (see e.g., Gitlin et al., 1999; Williams and Coles, 2007; Thomm et al., 2021b). Thereby, different approaches exist that attempt to systematize these different barriers by using categorizations that range from rather broad distinctions between knowledge- and motivation-related barriers (e.g., Kiemer and Kollar, 2021) to more nuanced ones (e.g., van Schaik et al., 2018; Brown et al., 2022). While van Schaik et al. (2018), for example, differentiate between research knowledge level, individual teacher level, school organizational level, and communication level to systematize barriers, Brown et al. (2022) use the categories benefit, cost, and signification.

Specifically referring to knowledge-related barriers, we argue that these are often associated with the structure of modern societies, more precisely with the cognitive division of labor (Bromme et al., 2015). In general, the cognitive division of labor is defined as the uneven distribution and use of knowledge due to training of highly specialized experts (Bromme et al., 2010). This leads to the fact that teachers, in most countries, are trained as experts in education and learning, but not as educational scientists. Hence, student teachers, for example, have lower abilities in science-related areas compared to other students (Besa et al., 2020; Thiem et al., 2020) and are hardly trained in informing practice with scientific evidence (Ostinelli, 2009). Consequently, they often lack educational research literacy (Shank and Brown, 2007), making so-called first-hand evaluation of scientific evidence (i.e., assessing the veracity of scientific information by relying on objective criteria such as evaluating the study design) challenging for them (e.g., Bromme et al., 2010, 2015; Bromme and Goldman, 2014; Hendriks et al., 2015; Brown et al., 2017).

As a result, teachers often rely on so-called *second-hand evaluations*, which are defined as the assessment of the credibility and trustworthiness of the information's source (e.g., Bromme et al., 2010; Bromme and Goldman, 2014; Merk and Rosman, 2019). Hence, instead of analyzing whether the information itself is "true" or not, they evaluate if they can trust the information at hand or the person or body who conveys it (e.g., the researcher or specific science communication formats like clearing houses). Consequently, teachers' trust in educational science can be seen as a central predictor to which extent teachers positively evaluate and thus engage with scientific evidence (Hendriks et al., 2016;

Bromme et al., 2022). Hence, in the following, we analyze what factors influence teachers' trust in educational science. More specifically, we investigate whether teachers' trust in educational information is influenced by the source of information, their trust in educational science and in science in general, as well as by their prior topic-specific beliefs in the sense of a confirmation bias. While the first research question aims at aligning the study at hand with previous research on teachers' trust in science, the latter two extend existing research by transferring findings on teachers' beliefs from the domains of epistemic beliefs and cognitive biases to research on trust in science. This is to find out whether teachers' beliefs influence teachers' trust in educational science which, in turn, could also influence teachers' engagement with scientific evidence.

2. Trust-related barriers to evidence-informed practice

There are several definitions of trust (see Dietz and Den Hartog, 2006 for an overview), which often share the common overarching idea that there is a one-way dependency in which one party (the so-called trustor) trusts another party (the so-called trustee; Mayer et al., 1995; Dietz and Den Hartog, 2006; Blöbaum, 2016). Thereby, according to the popular definition of Mayer et al. (1995), the trustor agrees "*to be vulnerable to the actions of (...) [the other] party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party*" (p. 712).

Furthermore, there are different conceptualizations of trust such as trust as a belief, trust as an action, or trust as a decision (see again Dietz and Den Hartog, 2006 for an overview). In the following, we primarily focus on *trust as a belief*, i.e., the evaluation of trustworthiness of the potential trustee (Dietz and Den Hartog, 2006), which is why the question arises as to what causes trustworthiness in the first place. In this regard, three dimensions of trust are typically mentioned: (1) expertise, (2) integrity, and (3) benevolence. A party seems trustworthy if he/she is having high and relevant knowledge on the topic of interest (expertise), adheres to the rules, norms, and values of his or her profession (integrity) without ignoring the interest in the good of others (benevolence, e.g., Mayer et al., 1995; Dietz and Den Hartog, 2006; Hendriks et al., 2016; Bromme et al., 2022). These dimensions can be ascribed to characteristics of the trustee (Mayer et al., 1995; Dietz and Den Hartog, 2006; Blöbaum, 2016), whereby it depends on the trustors' perception how strong these characteristics are (Blöbaum, 2016). Characteristics of the trustor him- or herself, such as generalized trust in institutions or other beliefs, can also influence the degree of trusting a party (Mayer et al., 1995; Dietz and Den Hartog, 2006; Blöbaum, 2016).

Based on this, we argue that evaluating trustworthiness is accompanied by subjective and, therefore, less generalizable criteria, which could be prone to errors based on trustors' belief systems. Hence, we focus on three characteristics of teachers as trustors—(1) trust in different sources, (2) trust in educational science and science in general, as well as (3) prior individual beliefs about specific educational topics. While the former two are more directly related to second-hand evaluations, we argue that the latter (prior beliefs) are also important in this context since they might influence teachers' trust in information provided by educational science and therefore could act as (rather indirect) barriers to evidence-informed practice.

2.1. Trust in different knowledge sources

As outlined in the introduction, teachers can use a variety of sources to inform their practice. This ranges from anecdotal evidence (also called experiential sources, Bråten and Ferguson, 2015) that includes empirical information based on, for example, own personal experiences or experiences from a colleague (Buehl and Fives, 2009; Kiemer and Kollar, 2021) to more formalized sources such as lectures and formalized bodies of knowledge like scientific evidence (i.e., research findings, Buehl and Fives, 2009). The presentation or communication form of information is independent of the source. For example, anecdotal but also scientific evidence can be communicated both orally and in writing. Previous studies show that (future) teachers name and recognize the variety of information sources themselves (Buehl and Fives, 2009), but when it comes to practical decisions, they prefer anecdotal over scientific evidence (e.g., Gitlin et al., 1999; Parr and Timperley, 2008; Buehl and Fives, 2009; Cramer, 2013; Bråten and Ferguson, 2015; Zeuch and Souvignier, 2016; Menz et al., 2021; Groß Ophoff and Cramer, 2022). Furthermore, the preference of anecdotal evidence does not only seem to be associated with student teachers' motivation to learn in teacher training (Bråten and Ferguson, 2015), but it was also identified as the root of their beliefs and persistent misconceptions about specific topics from educational psychology (Menz et al., 2021).

In addition, the predominance of anecdotal evidence among (future) teachers is also evident in studies on trust in different sources: For example, Landrum et al. (2002) focused on an overall assessment of trust among student teachers by comparing, among others, the sources "scientific journals" vs. "teachers." In this descriptive study, the participants trusted information provided by other teachers more than information provided by scientific journals. In a two-step— first exploratory and then confirmatory—study by Merk and Rosman (2019) the results were more differentiated as student teachers deemed scientists in educational science as "smart but evil" (p. 6). Accordingly, they attributed more expertise but less benevolence and integrity to scientists than to practicing teachers. In a follow-up study, Rosman and Merk (2021) examined teachers' reasons for (dis-)trusting educational science vs. science in general. In line with the "smart but evil" pattern outlined above, teachers more

strongly emphasized integrity and benevolence—compared to expertise—as reasons for distrusting educational scientists. Similarly, Hendriks et al. (2021) found differences in student teachers' reasons for (dis-)trusting information by educational psychology scientists or teachers. In their descriptive study, student teachers deemed educational psychology scientists not only less benevolent but also as having less expertise than teachers. When specifically looking for practical advice, student teachers rated teachers as more trustworthy than scientists—consistent across all three dimensions (expertise, integrity, benevolence).

To sum up, these results suggest that different knowledge sources influence (future) teachers' second-hand evaluation in a manner that anecdotal evidence provided by practicing teachers is perceived as more trustworthy than scientific evidence, leading to greater use of anecdotal evidence in practice. We argue that this preference may not only be caused by differences in the epistemological nature of anecdotal evidence (i.e., a non-scientific and possibly more "user-friendly" body of knowledge) compared to scientific evidence (which is often more abstract and theoretical) but also by the fact that individuals from one's own in-group are often evaluated more positively compared to individuals from out-groups since they share the same profession-related experiences (e.g., Mullen et al., 1992). Hence, educators who share anecdotal evidence from their day-to-day practice might seem more trustworthy to teachers compared to scientists. The preference for anecdotal evidence coming from other teachers could then act as a barrier to evidence-informed school practice, and, in the worst case, as Rosman and Merk (2021) argue, lead to dysfunctional practices when decisions are, for example, built on passed-on misconceptions like the prominent so-called neuromyths that can appear at an early career stage and are thus already common among student teachers (Krammer et al., 2019, 2021). There are many neuromyths with the learning style myth as a well-known example that has been debunked years ago (e.g., Pashler et al., 2008), but is still very popular among practitioners (Krammer et al., 2019, 2021). If (student) teachers do not consult scientific evidence, in light of this myth "they may waste time developing teaching materials tailored to individual students' learning styles" (Rosman and Merk, 2021, p. 1). Hence, we will analyze whether teachers trust knowledge claims by other teachers and scientific studies differently. To test this research question, we formulate the following first hypothesis:

> H1: *When teachers are confronted with knowledge claims regarding specific topics from educational science, they show more trust in claims if these are allegedly from another teacher (anecdotal evidence) than from a scientific study (scientific evidence).*

2.2. Trust in (educational) science

In the context of teachers' evaluation of scientific information, epistemic beliefs also play a pivotal role (Bendixen and Feucht, 2010; Fives and Buehl, 2010). Epistemic beliefs are defined as "individuals' beliefs about the nature of knowledge and the process of knowing" (Muis et al., 2016, p. 331). These beliefs can be simultaneously domain-general and domain-specific, whereby beliefs relating to different domains can influence each other (e.g., Buehl and Alexander, 2001; Muis, 2004). Hence, Muis et al. (2006) proposed the Theory of Integrated Domains in Epistemology (TIDE) that refers to the interplay of general epistemic beliefs, academic epistemic beliefs as well as domain-specific epistemic beliefs, which, in turn, are influenced by different contextual factors (i.e., the socio-cultural, academic, and instructional context). In their framework, the authors define general epistemic beliefs as "beliefs about knowledge and knowing that develop in nonacademic contexts such as the home environment, in interactions with peers, in work-related environments, and in any other nonacademic environments" (Muis et al., 2006, p. 33). Academic epistemic beliefs, on the other hand, encompass "beliefs about knowledge and knowing that begin to develop once individuals enter an educational system" (Muis et al., 2006, p. 35). Furthermore, these two belief dimensions can be differentiated from domain-specific epistemic beliefs—"beliefs about knowledge and knowing that can be articulated in reference to any domain to which students have been exposed" (Muis et al., 2006, p. 36). In 2018, Merk et al. (2018) extended the framework by adding topic-specific beliefs, i.e., beliefs regarding specific topics or theories, as a further dimension of epistemic beliefs. By analyzing student teachers' epistemic beliefs according to different educational topics, they found (at least on a correlational level) empirical support for the predictions of their framework that (1) topic-specificity is a feature of epistemic beliefs and that (2) domain-specific beliefs and topic-specific beliefs influence each other reciprocally.

Transferring the predictions of the TIDE framework to teachers' evaluation of trustworthiness, we posit that low trust in educational science and in science in general could be a barrier to evidence-informed practice via its effects on topic-specific trust. In fact, following the framework's assumption of different levels of beliefs reciprocally influencing each other, scientific findings about specific educational topics would also be deemed as less trustworthy and, therefore, less relevant for teaching practice in teachers with low trust in educational science and in science in general. In other words, we investigate whether there is an interplay between teachers' trust in specific knowledge claims from educational science, their trust in educational science, and their global trust in science. More specifically, we formulate the following second hypothesis:

H2: *Trust in specific knowledge claims from educational science can be predicted by (domain-specific) trust in educational science and global trust in science.*

2.3. Confirmation bias

The last factor considered in the present article involves teachers' prior beliefs about the corresponding knowledge claims. An extensive body of studies has shown that individuals' prior beliefs on a specific topic (e.g., on psychological or political issues) influence the search for and interpretation of information on this topic. In the literature, this phenomenon is labeled with different terminologies such as prior attitude effect (Druckman and McGrath, 2019), biased assimilation (Lord et al., 1979; Lord et al., 1984), or congeniality bias (Hart et al., 2009), but the most prominent one is *confirmation bias* (e.g., Lord et al., 1979; Nickerson, 1998; Oswald and Grosjean, 2004). Confirmation bias can be divided into two subcomponents: selective exposure and selective judgment. Selective exposure comes into play while seeking information on the respective topic, and manifests itself in preferring belief-consistent and ignoring belief-inconsistent information (Hart et al., 2009; Stroud, 2017). Selective judgment, on the other hand, is defined as the process of interpreting information in a way that—irrespective of the veracity of that information—this information is preferred if it is consistent with one's prior belief. By contrast, if this information is in conflict with one's prior belief, it is either quickly discounted or analyzed thoroughly to identify errors in it (e.g., Lord et al., 1979; Nickerson, 1998; Jonas et al., 2001; Oswald and Grosjean, 2004; Stroud, 2017). However, these two subcomponents are not always clearly differentiated and sometimes confirmation bias is referred to even when only one of the two subcomponents is considered (e.g., Butzer, 2020).

If teachers are subject to confirmation bias, this can act as a barrier to evidence-informed practice: On the one hand, teachers may completely distrust scientific evidence on a specific educational topic if the evidence is contradictory to their previous beliefs, resulting in ignoring the scientific evidence in their practical actions. On the other hand, they may selectively trust scientific evidence that is in line with their beliefs, and thus, selectively use scientific evidence. Even though confirmation bias is already discussed as a barrier to evidence-informed practice (e.g., Katz and Dack, 2014; Andersen, 2020), it has been less systematically analyzed in this context so far. While there are a few qualitative studies that draw attention to the existence of confirmation bias among teachers in the context of data-based decision making (e.g., Van Lommel et al., 2017; Andersen, 2020), only Masnick and Zimmerman (2009) have explicitly analyzed whether confirmation bias—or more precisely selective judgment—influences, among others, student teachers' evaluation of scientific evidence. In line with

selective judgment, participants perceived the arrangement of the study as more appropriate as well as the results as more important and interesting when the results were in line with their beliefs compared to when they contradicted them.

It should be noted that Masnick and Zimmerman (2009)'s dependent variables conformed to a first-hand evaluation of the study in question (e.g., evaluating the appropriateness of a study's design). However, despite the importance of second-hand evaluation in teachers' dealing with scientific evidence (see section 1), the influence of confirmation bias on such second-hand evaluations has, to the best of our knowledge, not been investigated so far. Consequently, we focus on the subcomponent selective judgment by examining whether teachers' prior topic-specific beliefs influence trust ratings in the sense of a confirmation bias. In doing so, we test the last hypothesis:

H3: *When teachers are confronted with evidence for specific knowledge claims of educational science, they show more trust in these claims if these are belief-consistent.*

3. Materials and methods

3.1. Experimental design

All hypotheses were preregistered (Schmidt et al., 2022). To test the hypotheses, we designed a 2 × 2 within-person experiment with the two independent variables *source of evidence* (scientific evidence from a published scientific study vs. anecdotal evidence from another teacher on a teacher blog) and *belief-consistency of evidence* (belief-consistent vs. belief-inconsistent claims). We thereby constructed texts about four topics from educational science (effects of retention, gender differences in grades, text-picture integration, signaling) for each source of evidence (i.e., scientific vs. anecdotal evidence), and with two variations of the claims made in texts (e.g., integrating text into pictures positively affects learning vs. does not positively affect learning) to allow us to manipulate belief-consistency. This resulted in 4 * 2 * 2 = 16 texts in total. These texts were mostly comparable in structure and wording and included a specific knowledge claim regarding the respective educational topic. For each topic, the only information that varied in the texts were the source of evidence and the belief-consistency of the presented knowledge claim (see Figure 1 for epitomes). To prevent respondent fatigue (Lavrakas, 2008) and unintentional unblinding, the participants, however, received only two out of these 16 texts (see below for details). All texts and all other study materials are publicly available (Schmidt et al., 2022).

Topic 1: Retention ("Repeating classes")

A subject of much debate among educational practitioners and scholars is, for example, whether retention helps weak students catch up.

To what extent do the following statements apply to you?

	rather agree	rather disagree	don't know / not specified
I think that class repetitions are actually beneficial for weak students.	☐	☐	☐

--next page--

On this topic, for example, Quantz & Peters (2014) write in the "Journal of Effective Teaching": In an educational science study with N = 1134 students, we found that class repetition is found to be associated with smaller achievement gains (on average) for weak students than promoting these students.

Topic 2: Integration of figures and texts

A subject of much debate among educational practitioners and scholars is, for example, whether it is beneficial to integrate explanatory texts in figures (rather than putting them in separate legends).

To what extent do the following statements apply to you?

	rather agree	rather disagree	don't know / not specified
I think that class repetitions are actually beneficial to integrate explanatory texts into figures.	☐	☐	☐

--next page--

On this topic, for example, the user "MrMueller" writes on the Teacher-Blog "HeartAndSoulTeacher.org": According to my experience over many years, I found, that students learn more successfully from materials in which figures and explanatory texts are integrated (compared to materials in which illustration and explanatory text are presented separately).

FIGURE 1
Epitomes of the 16 texts. Italic: Same for all sources, beliefs, and topics; Underscored: Same for both sources and beliefs within each topic; Bold: Same for every topic and both beliefs within both sources; Gray: Same for both sources within every topic and belief combination.

3.2. Procedure and measurements

At the beginning of the experiment, trust in science and trust in educational science (independent variable $H2$) were measured by self-assessment using items from the science barometer (Weißkopf et al., 2019). This large trend study uses the following item to assess trust in science: *"How much do you trust in science and research?"* along with a five-point Likert-scale (1 = *trust completely* to 5 = *don't trust*; with a *don't-know/not specified*-option). Additionally, we used the same item stem to assess trust in "educational science and educational research" (independent variable $H2$). To ensure that all participants conceptualize educational science and educational research in a similar way, we provided a brief explanation of educational science and educational research by defining it as an area *"[...] that deals with the theory and practice of education."* We further provided examples of subdisciplines of educational science and educational research, which *"are, among others, educational science, educational psychology, economics*

of education as well as sociology of education." Subsequently, participants were introduced to a first (randomly chosen) topic (e.g., text-picture integration), and their belief toward this topic (auxiliary variable to construct the independent variable for $H3$) was assessed using a trichotomous item (e.g., *"I think it is actually beneficial to integrate text into pictures"*; with the answer options *"rather agree," "rather disagree,"* and *"don't know/not specified"*). Thereafter, they were presented with the randomly chosen text that included a knowledge claim referring to this topic using evidence from a randomly chosen source (independent variable $H1$; e.g., scientific evidence from a published scientific study) and of randomly chosen belief-consistency (e.g., consistent; independent variable $H3$). Referring to the latter, the knowledge claim presented was dependent on the previously stated belief of the participants and thus not *per se* an inconsistent or consistent claim (e.g., if the participant believes that integrating text into pictures is beneficial, the consistent knowledge claim informs about the benefit of integrating text into pictures). Finally, respondents were prompted to rate their trust in the respective knowledge claim (dependent variable for all three hypotheses) using an adapted item from the science barometer (Bromme et al., 2022; e.g., *"How much do you trust the claims of the educational scientists Quantz & Peters on the topic of text-picture integration?"*) on a five-point Likert-scale again ranging from 1 = *trust completely* to 5 = *don't trust* with a *don't-know/not specified*-option. Subsequently, this procedure was repeated for the second randomly chosen topic. Thereby, participants were subjected to—compared to the first round—the opposite condition of source and belief-consistency of evidence. For example, if scientific evidence including a consistent knowledge claim was presented in the first round (by random drawing, see above), the second knowledge claim referred to anecdotal evidence that was contradictory to the participants stated belief.

3.3. Sample

To achieve external validity, we commissioned a service provider to recruit a representative sample of in-service teachers in Germany. This was achieved using both random digit dialing (Wolter et al., 2009) and *post hoc* inverse-probability weighting (Mansournia and Altman, 2016). To avoid Type II errors, we conducted power analyses assuming random intercept regression models with level-1 dummy variables ($H1$ and $H3$) or a continuous level-2 predictor ($H2$) and small to moderate effect sizes (see Preregistration, Schmidt et al., 2022). Specifying a sample size of $N = 400$ participants, this resulted in good power estimates (>90%) for both model types. Correspondingly, the field provider stopped sending invitations and reminders after a sample size of $N = 400$ was reached, thus resulting in $N = 414$ participants. The distribution of our weighted sample and the corresponding population are given in Table 1.

	Population	Empirical
School type		
Elementary School (the German Grundschule)	29.3	29.3
Lower Secondary School (the German Hauptschule)	4.1	4.1
Intermediate Secondary School (the German Realschule)	8.0	8.0
Upper Secondary School (the German Gymnasium)	25.8	25.8
Comprehensive School (the German Gesamtschule)	19.9	20.0
Special School (the German Förderschule)	10.0	10.0
Age		
Age < 30	7.3	7.4
$30 \leq$ Age < 59	80.0	80.8
$60 \leq$ Age	12.4	11.8
Gender		
Men	26.9	26.9
Women	73.1	73.1

3.4. Statistical analysis

As preregistered, we used multi-level linear models (Gelman and Hill, 2007) to investigate the effects of the two experimentally manipulated independent variables (*source of evidence* and the *belief-consistency of evidence*; see Schmidt et al., 2022). As we did not preregister a detailed analysis plan, we did not include information on how to handle missing values. However, as our data contained a nontrivial amount of missing data after recoding the answer options "don't know/not specified" as missings (Lüdtke et al., 2007; 2% overall and 15% in the dependent variable), we decided to multiply impute these missing values using chained equations (van Buuren, 2018) and handle the multi-level structure of the data within these imputations using a dummy indicator approach (Lüdtke et al., 2017).

After carefully checking the imputation chains and the distributions of the imputed values, we estimated Bayesian random intercept models with flat priors for the regressions weights using the R package brms (Bürkner, 2017), which is based on the probabilistic programming language Stan (Stan Development Team, 2017). This package allows to incorporate survey weights by different contributions of data points to the likelihood, and also has built-in capacity for dealing with multiply imputed data: Distinct models are fitted for each imputed data set, resulting in as many models as imputations. While combining these models (model pooling) is a complex

task in classical statistics (Rubin, 1976), it is straightforward after Bayesian estimation: One has just to join the posterior draws of the submodels (Zhou and Reiter, 2010). Furthermore, for a better interpretation, we have standardized the dependent variables and continuous predictors.

To evaluate not only the predictors but also the whole regression models, we estimated *Conditional R^2* (Gelman et al., 2019) and compared, in cases where highest density intervals (HDIs) of the predictors indicate evidence for the null-hypothesis or negligible to small effects, the predictive performance of the models using Bayes factors based on bridge sampling (Gronau et al., 2017).

4. Results

All detailed results can be retrieved from the publicly available Reproducible Documentation of Analysis (Schmidt et al., 2022). To gain first insights into the results of our experiment, we plotted weighted means and standard deviations of the dependent variable trust by source of evidence, belief-consistency of evidence, and topic in Figure 2. These descriptive statistics imply (descriptive) evidence against $H1$, namely that teachers show more trust in claims regarding specific topics if these are allegedly from another teacher (anecdotal evidence) than from a scientific study (scientific evidence). In fact, all means of the trust variables were higher for the scientific study source, regardless of the respective topic and belief-consistency combination. Overall, this effect showed a large Cohen's $d = -0.81$, which varied substantially over the topics (gender differences in grades: $d = -1.25$, text-picture integration: $d = -1.04$, effects on retention: $d = -0.26$, signaling: $d = -0.86$).

Hypothesis 2 can be evaluated descriptively using Figure 3. As the bubble sizes in Figure 3 are proportional to the number of observations for each plot, one can see that there are some relations between trust in knowledge claims from educational science and trust in educational science (domain specific trust) as well as between the latter and trust in science (global trust). In contrast, trust in science seems to be uncorrelated with trust in knowledge claims from educational science. When only educational knowledge claims stemming from scientists, i.e., scientific evidence, are considered trust in science seems to be correlated with trust in knowledge claims from educational science, too. Furthermore, this visual impression is also reflected by descriptive correlation measures (see Pearson's r and Kendall's τ in Figure 3).

Finally, Figure 2 points toward a verification of Hypothesis 3, which posits that teachers trust claims from educational science more if these are belief-consistent. In fact, for each topic and source combination, participants consistently reported greater trust in the source if the respective claim was belief-consistent (overall: $d = 0.40$, gender differences in grades: $d = 0.43$, text-picture integration: $d = 0.28$, effects on retention:

$d = 0.64$, signaling: $d = 0.25$). Details on participants' beliefs about each topic can be found in Table 2.

To back up these descriptive results with inferential statistics, we estimated a series of Bayesian random intercept models. We started with a model including only the random

FIGURE 2
Weighted means and standard deviations of trust by belief-consistency of evidence, source of evidence, and topic.

intercept and dummy variables indicating the topic of the texts (see Table 3). These two predictors resulted in a $Conditional\ R^2 = 0.23$, whereas the addition of a dichotomous coded variable indicating the source (Model 2; referring to $H1$) resulted in an unstandardized coefficient of $b = 0.59$ (95%HDI: [0.48; 0.70]) and an increase of $Conditional\ R^2$ to 0.38. In the next two models, we consecutively included the continuous predictors trust in educational science and trust in science to test $H2$. The HDI of the standardized slope of trust in educational science did not contain zero (Model 3), whereas the HDI of the standardized slope for trust in science did so (Model 4). Both predictors explained comparatively little additional variance ($Conditional\ R^2 = 0.39$ for Model 3 and $Conditional\ R^2 = 0.39$ for Model 4). Exploratory (deviating from the preregistration) computed Bayes factors comparing Model 2 with Model 3 and Model 3 with Model 4 revealed some further evidence for this interpretation ($BF_{23} = 0.001$ and $BF_{34} = 4.86$). To test whether the findings in Model 4 on trust in science are potentially confounded by trust in educational science, we specified an additional model (Model 3a) that only included trust in science as a predictor. This exploratory (deviating from the preregistation) analysis revealed that including only trust in science as a predictor again explained comparatively little additional variance ($Conditional\ R^2 = 0.39$). The HDI of the standardized coefficient for trust in science was narrow (95%HDI: [0.00; 0.16]), which can be interpreted as evidence for a negligible effect. In Model 5, we included the third experimentally varied dichotomous variable, belief-consistency, as a dichotomous indicator (referring to $H3$). The

FIGURE 3
Associations of trust in knowledge claims from educational science respectively scientific evidence, educational science and science in general. Bubble plots. The expression *trust in knowledge claims from educational science* refers to trust in the knowledge claims across both sources of evidence (anecdotal and scientific evidence), whereas *trust in scientific evidence* only refers to the source scientific evidence. The bubble areas are proportional to the number of observation for each subplot. r, Pearson's r; τ, Kendall's τ.

corresponding unstandardized regression coefficient was $b = -0.35$ (95%HDI: $[-0.47; -0.23]$) and *Conditional* R^2 increased again to 0.41. In Model 5a, we analyzed the effect of belief-consistency specifically on scientific evidence, which resulted in an unstandardized regression coefficient of $b = -0.38$ (95%HDI: $[-0.55; -0.21]$) and an *Conditional* R^2 of 0.91. To further explore (a further deviation from the preregistration) the influence of trust in educational science, trust in science, and belief-consistency specifically on scientific evidence, we estimated an interaction model (Model 6) to see whether the source of evidence moderates the association between trust in educational science, respectively, trust in science and trust in knowledge claims from educational science. The HDI of the standardized coefficient of the interaction term of trust in educational science and source ($\beta = 0.11$; 95%HDI: $[-0.00; 0.22]$) as well as the HDI of the standardized coefficient for the interaction term of trust in science and source of evidence ($\beta = 0.11$; 95%HDI: $[0.00; 0.22]$) were adjacent to zero which can be interpreted as evidence for a small moderation effect of the source. However, Bayes factors comparing Model 5 with Model 6 revealed evidence in favor of Model 5 ($BF_{56} = 0.33$).

5. Discussion

In the present study, we focused on trust-related barriers teachers might face when it comes to realizing evidence-informed practice. As teachers often evaluate scientific evidence from educational science by evaluating whether the person or body who conveys the evidence appears trustworthy, we analyzed factors that might influence such second-hand evaluations. We thereby examined (1) whether teachers trust knowledge claims by other teachers and scientists differently, (2) whether there is an interplay between teachers' trust in specific knowledge claims from educational science, their trust in educational science, and their global trust in science, as well as (3) whether teachers' prior topic-specific beliefs influence their trust ratings in the sense of a confirmation bias.

5.1. Summary and discussion of main results

A key finding of our study is that teachers consider knowledge claims made by educational scientists (on average) as more trustworthy than those made by other teachers. This result is particularly surprising because it contradicts numerous previous research results on the preference of anecdotal over scientific evidence (e.g., Landrum et al., 2002; Merk and Rosman, 2019; Hendriks et al., 2021; Rosman and Merk, 2021). In the following, we want to discuss possible reasons for this deviation:

One explanation could be attributed to the fact that previous studies mainly used student teacher samples (Landrum et al., 2002; Merk and Rosman, 2019; Hendriks et al., 2021). As we have surveyed in-service teachers, the deviation might be (partially) due to the different samples' characteristics such as differences in working or practical experiences. By providing evidence that college experiences influence student teachers' beliefs about knowledge sources, a study by Perry (1999) supports this idea.

Differences in the operationalizations of dependent and independent variables between previous studies and our study can offer further explanations. Referring to the dependent variable "trust," previous studies primarily measured dimensions of trust (expertise, integrity, and benevolence) rather than an overall assessment of (dis-)trust in science. In these studies, trust ratings regarding science and scientists in general were lower than those regarding teachers—especially in terms of integrity and benevolence—but globally still rather high (Merk and Rosman, 2019; Hendriks et al., 2021; Rosman and Merk, 2021). The operationalization of the independent variable "source" also deviates from previous operationalizations. While previous studies focused on trust in the source itself without referring to any content-related information (Landrum et al., 2002; Hendriks et al., 2021; Rosman and Merk, 2021), we focused on trust in specific *claims made by a specific source*. The inconsistencies of our findings regarding prior research might thus be caused by the fact that we kept all knowledge claims mostly constant over the different sources, whereas, in other studies, there is a potential confounding between source and content. Asking participants about different sources without referring to specific topics may introduce bias since certain knowledge sources are usually associated with specific topics. To give an example, if teachers are asked about their trust in the expertise of teaching practitioners, they might think of expertise in classroom teaching, whereas, when asked the same question with reference to educational scientists, they might consider researchers' expertise in conducting empirical studies (Merk and Rosman, 2019). In this case, there is thus no consistent criterion for comparison. Hence, the finding, in previous studies, that teachers are often reluctant to use scientific knowledge for their day-to-day practice may not result from the fact that they generally doubt that science is able to deliver robust and trustworthy knowledge, but possibly due to the fact that scientific evidence is often associated with more abstract and theoretical information (e.g., Buehl and Fives, 2009; Bråten and Ferguson, 2015; Groß Ophoff and Cramer, 2022).

Furthermore, our results provide early evidence against an influence of an in-group bias (e.g., Mullen et al., 1992) on teachers' trust. The knowledge claims provided by other teachers, i.e., by an in-group member, were not perceived as more trustworthy than those by scientists—out-group members. However, this statement should be taken with caution. Although we kept the knowledge claims and wording mostly constant, the source "teacher" referred to some other (unknown) teacher who published own experiences on a teacher blog and the source "scientists" to scientists who published their findings

TABLE 2   Beliefs per topic in absolute frequencies.

Topic: Statement	Frequencies		
	Rather agree	**Rather disagree**	**Don't know/Not specified**
Gender differences in grades: "I think girls actually get better grades."	92	107	12
Text-picture-integration: "I think it is actually beneficial to integrate text into pictures."	139	39	27
Effects of retention: "I think class retention is actually beneficial for weak students."	151	58	11
Signaling: "I think it is actually beneficial to signal central information in texts."	138	43	11

in a scientific journal. Hence, the publication body (teacher blog vs. scientific journal) might also have influenced teachers' trust ratings. As Bråten and Ferguson (2015) found that student teachers believed least in knowledge stemming from social and popular media compared to anecdotal and formalized knowledge, blog posts, as a form of social media, could have decreased teachers' trust in the provided claims by other teachers. One possible explanation for the lower trust in anecdotal evidence on teacher blogs could be that anyone could write such posts, which can reduce its seriousness and therefore its trust in it. Another factor that might influence teachers' trust ratings in anecdotal evidence could be related to the familiarity of the source. Teachers might perceive a claim from a trusted colleague as more trustworthy compared to an anonymous blog-poster. Given that most previous studies investigating (student) teachers' trust in different sources also did not explicitly take source familiarity into account, but rather used generic terms such as an "experienced teacher" (Landrum et al., 2002, p. 44), a "practitioner" (Merk and Rosman, 2019, p. 4), or "a teacher who has taught at a school for a number of years" (Hendriks et al., 2021, p. 170), source familiarity might be an interesting independent variable for further research.

All in all, however, this finding is good news as a general lack of trust in statements from educational science does not seem to be a barrier to evidence-informed action in schools. But at the same time, and in line with previous research, we found strong support for a substantial confirmation bias in teachers' trust ratings. Confirmation bias cannot only come into play while evaluating scientific information first-hand (e.g., Masnick and Zimmerman, 2009), but can also distort, as demonstrated in the present study, teachers' second-hand evaluations. Thus, it can be assumed that teachers evaluate knowledge claims to a large extent in such a way that they confirm their own prior beliefs. Consequently, they might be highly selective in choosing the evidence they refer to. The general finding that teachers have traditionally been rather reluctant to turn to scientific evidence and rely heavily on their professional autonomy in making decisions (e.g., Landrum et al., 2002; Buehl and Fives, 2009; Bråten and Ferguson, 2015; Groß Ophoff and Cramer, 2022) could, thus, be viewed in a differentiated way: It is not a general lack of trust in claims from educational science that might hinder teachers from engaging with scientific evidence,

but the question of which filters come into play to evaluate the evidence. The present study revealed that confirmation bias might work as one such filter and, thus, as a potential barrier for evidence-informed practice as teachers might not easily change their practice in light of scientific evidence that does not fit their beliefs. Given that many students already enter teacher education with a specific set of misconceptions such as neuromyths (Krammer et al., 2019, 2021), our findings on the existence of a confirmation bias are particularly worrying, since confirmation bias may lead to a further strengthening of such misconceptions. Furthermore, if teachers trust evidence that is consistent with their beliefs more and rather distrust evidence that is contradictory, they will continue to teach as before. This seems rather unproblematic as long as the teachers' practice is tried and tested. In addition, it is quite unrealistic to expect teachers to inform every practical decision and action with evidence. However, it is, on the one hand, problematic when it comes to concepts that are scientifically untenable but continue to persist in practice, and, on the other hand, when evidence is used to develop new approaches or to overcome hitherto unsolved problems as this makes it difficult to stimulate new avenues. Therefore, approaches must be identified to specifically motivate teachers to change their practice when scientific evidence contradicts their beliefs and existing practice.

Furthermore, of course not all teachers trust claims from educational science in general equally. This can be explained partially by referring to some of the predictions of the TIDE framework (Muis et al., 2006). Admittedly, we found evidence for a negligible influence of global trust in science on topic-specific trust, a finding that might be caused by different associations between science and educational science. For example, the former might be associated with the mixing of different chemicals in a laboratory, whereas the latter might be perceived more as a process of developing abstract theories. Nevertheless, in line with the TIDE-framework, domain-specific trust in educational science predicted teachers' topic specific trust. As a consequence, those teachers who perceive educational science in general as less trustworthy also report less trust in specific claims from educational science, and thus might ignore scientific evidence as a source of information for their practical actions. Therefore, when striving to foster teachers'

TABLE 3 Results of the Bayesian multi-level models.

| | Trust | | | | | | | | | | | | | | | |
| | Model 1 | | Model 2 | | Model 3 | | Model 3a | | Model 4 | | Model 5 | | Model 5a | | Model 6 | |
	Est.	HDI (95%)	Est.	HDI (95%)	Est.	HDI (95%)	Est.	HDI (95%)	Est.	HDI (95%)	Est.	HDI (95%)	Est.	HDI (95%)	Est.	HDI (95%)
Intercept	−0.15	−0.29 to −0.02	−0.44	−0.58 to −0.31	−0.45	−0.59 to −0.31	−0.44	−0.58 to −0.30	−0.45	−0.58 to −0.31	−0.27	−0.41 to −0.12	0.46	0.27 to 0.65	−0.27	−0.41 to −0.12
$I_{Topic}^{Integration}$	0.20	0.02 to 0.38	0.19	0.03 to 0.36	0.20	0.03 to 0.36	0.19	0.03 to 0.36	0.20	0.03 to 0.36	0.18	0.01 to 0.34	0.19	−0.05 to 0.43	0.17	0.01 to 0.34
$I_{Topic}^{Retention}$	0.24	0.06 to 0.42	0.21	0.04 to 0.37	0.21	0.05 to 0.37	0.20	0.04 to 0.37	0.21	0.04 to 0.37	0.20	0.04 to 0.36	−0.09	−0.32 to 0.14	0.20	0.04 to 0.36
$I_{Topic}^{Signaling}$	0.12	−0.06 to 0.31	0.15	−0.02 to 0.32	0.16	−0.01 to 0.33	0.15	−0.02 to 0.32	0.16	−0.01 to 0.33	0.16	−0.01 to 0.32	0.01	−0.23 to 0.26	0.16	−0.01 to 0.32
$I_{Source}^{ScientificEvidence}$			0.59	0.48 to 0.70	0.59	0.49 to 0.70	0.59	0.48 to 0.70	0.59	0.48 to 0.70	0.60	0.49 to 0.71			0.60	0.49 to 0.70
Trust in Educational Science					0.15	0.08 to 0.23			0.14	0.06 to 0.22	0.15	0.07 to 0.22	0.21	0.12 to 0.29	0.09	−0.00 to 0.19
Trust in Science							0.08	0.00 to 0.16	0.05	−0.03 to 0.12	0.04	−0.04 to 0.12	0.09	−0.00 to 0.17	−0.02	−0.11 to 0.08
$I_{Belief\text{-}Consistency}^{Inconsistent}$											−0.35	−0.47 to −0.23	−0.38	−0.55 to −0.21	−0.35	−0.47 to −0.22
Trust in Educational Science*$I_{Source}^{ScientificEvidence}$															0.11	−0.00 to 0.22
$I_{Source}^{ScientificEvidence}$*Trust in Science															0.11	0.00 to 0.22
Random effects																
σ^2	0.78		0.62		0.62		0.62		0.62		0.60		0.64		0.59	
τ_{00}	0.22		0.29		0.28		0.29		0.28		0.26		0.19		0.27	
ICC	0.22		0.32		0.31		0.32		0.31		0.30		0.77		0.31	
Clusters	412		412		411		412		411		411		411		411	
Observations	824		824		822		824		822		822		411		822	
Marginal R^2/ Conditional R^2	0.01/0.23		0.10/0.38		0.12/0.39		0.10/0.39		0.13/0.39		0.16/0.41		0.14/0.91		0.16/0.42	

The dependent variables [Trust in Knowledge Claims From Educational Science (Model 1–Model 5 and Model 6) and Trust in Scientific Evidence (Model 5a)] and continuous predictors (Trust in Educational Science, Trust in Science) are standardized. Est., estimate of the regression coefficient; HDI (95%), 95% highest density interval; σ^2, Level-1 residual variance; τ_{00}, Level-2 residual variance; ICC, intraclass correlation coefficient; Marginal R^2, variance that is explained by fixed factors; Conditional R^2, variance that is explained by fixed and random factors.

evidence-informed practice, one promising way may either focus on increasing trust in educational science as a whole, or on increasing trust in specific knowledge claims from that domain. However, so far little is know about how to actually do this. But even if an effective way to increase trust can be identified, we want to point out that trust in science is only a predictor in the sense of a necessary condition for acting in an evidence-informed manner, but it does not automatically imply evidence-informed actions. The same also applies to first-hand evaluation: Even a competent first-hand evaluation of the veracity of scientific evidence (e.g., enabled by comprehensive teacher training including a fundamental training in methodology and well-designed science communication) does not automatically imply engagement with evidence. This is also illustrated by the implementation steps of evidence-informed practice: after evaluating scientific evidence, teachers still need to link the evidence to their own prior knowledge and then need to find ways to concretely use the evidence in their practice (Brown et al., 2022). Nevertheless, it can be assumed that a successful first-hand evaluation supports correct receptions of scientific evidence, which in turn is a central basis for an adequate transfer of scientific evidence into practice.

In addition, we want to highlight that processes in schools and classrooms are complex, characterized by interpersonal interaction, and exposed to uncertainty in the field of educational action, which can be reduced but never completely resolved by recourse to scientific evidence (Cochran-Smith et al., 2014). Thus, even substantial engagement with evidence does not grasp the action situation in its whole complexity. In other words, the assumption of simple and fitting evidence for certain or even all conceivable questions of school practice is neither tenable nor scientifically justifiable (Renkl, 2022). Evidence in itself can also be problematic if low-quality evidence is referred to, such as findings that cannot be replicated (e.g., Makel and Plucker, 2014; Gough, 2021). Even within science, there are critical voices concerning the informativeness of common research methods (e.g., randomized controlled trials) for educational practice, which is, after all, highly context-specific (e.g., Berliner, 2002). However, we argue that understanding evidence-informed practice as an educational practice where scientific evidence is reinterpreted against the background of one's own experience and the context at hand (Brown et al., 2017) may counteract this criticism. Consequently, teachers do not only need to be able to evaluate, understand, and deal with scientific evidence, but also to reflect it in light of other information sources, for example, their own practical experiences, contextual knowledge, and local school data (e.g., Bauer et al., 2015; Brown et al., 2017). Only then can scientific evidence unfold its potential to contribute to a broadening of perspectives on the pedagogical field of action.

5.2. Methodological limitations and future research

Of course, our study is not without limitations. Some have already been mentioned before. In the following, we will briefly repeat these limitations and add further ones, as well as derive some implications for future studies.

The results of the present study indicate—and this is contrary to previous findings—that teachers perceive scientific evidence as more trustworthy than anecdotal evidence from other teachers. As mentioned in the section above, the operationalization of trust either as a multi-dimensional construct (with the dimensions expertise, integrity, and benevolence) or as an overall rating (e.g., used in the science barometer, Weißkopf et al., 2019) as well as the operationalization of the source (with or without content-related information) could explain the difference. With regard to the latter, we used claims regarding specific topics from educational science to increase internal validity. This approach avoids that participants make different, and thus not comparable, associations when thinking about the expertise of practicing teachers or educational scientists (see section above). In contrast, the external validity of our study might be curtailed through this approach, which is why our results should not simply be generalized to other educational science topics.

Furthermore, as outlined in the section above, the additional information about the publication body (teacher blog vs. scientific journal) may have confounded teachers' trust ratings in different sources. The same could also apply to the names of the persons who provide the information (Quantz and Peters, 2014 vs. Mr. Mueller) as well as the name of the publisher (Journal of Effective Teaching vs. HeartAndSoulTeacher.org), given that textual features of "scientificness" have been shown to affect the processing and evaluation of textual information (Thomm and Bromme, 2012). However, to increase the external validity of our material, we tried to create scenarios which are as realistic as possible. As teachers often do not meet scientists in person, we decided to use a written format of scientific evidence. To increase internal validity, we kept our study materials as parallel as possible across the two different sources (i.e., scientific and anecdotal evidence), which is why we also chose a written format for anecdotal evidence. In this context, blogs are typical for informing others about personal experiences in a written manner (e.g., Ray and Hocutt, 2006; Deng and Yuen, 2011). However, our approach reflects the challenge of designing internally valid study materials quite well: The more information is given in study materials, the less abstract the information appears and the more varying associations of the participants can be prevented. At the same time, it cannot be ruled out that one of these pieces of contextual information (e.g., publication body) confounds the variables that are actually relevant for the

study and introduces bias on the outcomes. Therefore, it might be reasonable to systematically vary the publication bodies in further studies.

Based on our findings, one might assume that teachers indeed have a high general trust in claims from educational science, which, however, might decrease when explicitly referring to concrete teaching practice. In other words, teachers might take the attitude that what scientists say is true and trustworthy, but has not much to do with their own teaching practice or with the issues they actually encounter in the classroom (e.g., Gitlin et al., 1999). This is a limitation of our study since we did not directly relate our study materials to school practice (or other contexts). However, we argue that these concerns are mitigated by the fact that our scientific and anecdotal evidence indirectly referred to school practice by presenting claims about an educational topic published in a teacher journal respectively on a teacher blog. Nevertheless, in future studies, the context in which the evidence is intended to be used could also be explicitly considered because Hendriks et al. (2021) found that this can lead to differences in trust ratings. In fact, in their study, student teachers perceived educational psychology researchers as more trustworthy than teachers when searching for theoretical explanations, but, in contrast, the source "teachers" was trusted more when it came to practical recommendations.

As a final limitation regarding our results on Hypothesis 1, we cannot completely exclude a Hawthorne effect or social desirability bias, meaning that trust in research could have been rated higher because participants were asked by researchers. Considering previous research, however, we see this influence as rather minor, as previous research studies have found greater trust in anecdotal evidence although their participants were asked by researchers, too.

Our results also show that selective trust in evidence may be fuelled by a confirmation bias. Recent studies illustrate that prior beliefs can even lead to the conclusion that when one is confronted with belief-inconsistent scientific evidence, certain topics cannot be scientifically investigated at all (Rosman et al., 2021; Thomm et al., 2021a). In addition, previous studies on confirmation bias indicate that the strength of prior beliefs moderates the influence of confirmation bias on searching and interpreting information (e.g., Taber and Lodge, 2006). Hence, in future studies, it would be reasonable to additionally collect data on the strength of teachers' prior beliefs, given that our results might underestimate the influence of confirmation bias on trust in teachers with strong prior beliefs and overestimate the influence in teachers with less entrenched prior beliefs. In addition to prior beliefs, other individual characteristics of teachers, but also characteristics of the scientific evidence can activate or act as filters (Fives and Buehl, 2012). With respect to our study, teachers' varying degrees of trust in the domain of educational science can be regarded as one such individual characteristic. However, factors like the individual

degree of educational research literacy or epistemic beliefs could be influential as well. Referring to the former, teachers with higher educational research literacy might generally indicate a higher trust in educational science, and possibly show less of a confirmation bias. With regard to epistemic beliefs, teachers with high multiplistic epistemic beliefs (scientific knowledge as subjective "opinions") might rate educational science as less trustworthy and be more inclined toward confirming their prior beliefs (Hofer and Pintrich, 1997). Referring to the characteristics of scientific evidence, future research should also examine (1) whether teachers trust certain research paradigms more than others (e.g., experimental vs. observational research), (2) if they trust scientific evidence more if it is proximal to their actual teaching practice, or (3) if they place higher trust into evidence that is close to their teaching subjects (e.g., because it is more familiar).

Finally, with regard to future research, we would like to emphasize that we have not measured teachers' engagement with evidence or use of evidence. Even though trust is a necessary (but not sufficient) predictor for engaging with evidence, it would be reasonable to additionally focus on actions based on trust (i.e., behavioral variables). In this regard, however, the conceptualization of objective measures that are intended to go beyond self-reporting (e.g., How likely is it that you will incorporate the research findings into your own practice?) is quite complex.

6. Conclusion

Taken together, our findings allow a more differentiated view of teachers' trust in educational science and, thus, of trust-related barriers teachers face when realizing evidence-informed practice: It is not a general lack of trust in science that might hinder teachers from engaging with (educational) scientific evidence, it is more about the filter function of beliefs that come into play to evaluate evidence that is problematic for an adequate realization of evidence-informed practice. Thereby, the present study revealed teachers' prior topic-specific beliefs as one such filter since they trusted more in evidence consistent with their prior beliefs than belief-inconsistent evidence (confirmation bias). Such selective trust can be dangerous since teachers may—if at all—inform their actions with empirical evidence, but just one-sidedly and not in its full scope. Moreover, we argue that second-hand evaluation as well as first-hand evaluation are necessary conditions to engage with evidence, which is why both needs to be fostered systematically. Nevertheless, a successful first-hand combined with a positive second-hand evaluation is an important foundation for evidence-informed practice but does not automatically imply evidence-informed actions and even if professional actions are informed by evidence, scientific evidence can only unfold its potential when it is reflected in light of information from other sources.

Data availability statement

Ethics statement

The study involving human participants was reviewed and approved by University of Tübingen Institutional Review Board (Protocol Number: A2.5.4-094_aa). The participants provided their written informed consent to participate in this study.

Author contributions

KS and SM: conceptualization, data curation, formal analysis, investigation, methodology, resources, software, supervision, validation, visualization, writing—original draft, review, and editing. TR: conceptualization, investigation, writing—original draft, review, and editing. CC: funding acquisition, writing—original draft, review, and editing. K-SB: writing—review and editing. All authors contributed to the article and approved the submitted version.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Andersen, I. G. (2020). What went wrong? Examining teachers' data use and instructional decision making through a bottom-up data intervention in Denmark. *Int. J. Educ. Res.* 102, 101585. doi: 10.1016/j.ijer.2020.101585

Bauer, J., and Prenzel, M. (2012). European teacher training reforms. *Science* 336, 1642–1643. doi: 10.1126/science.1218387

Bauer, J., Prenzel, M., and Renkl, A. (2015). Evidenzbasierte Praxis – im Lehrerberuf?! Einführung in den Thementeil (Evidence-based practice in teaching?! Introduction to the special issue). *Unterrichtswissenschaft* 43, 188–192. doi: 10.3262/UW1503188

Bendixen, L. D., and Feucht, F. C. (2010). "Personal epistemology in the classroom: What does research and theory tell us and where do we need to go next?," in *Personal Epistemology in the Classroom, 1st Edn.*, eds L. D. Bendixen and F. C. Feucht (Cambridge, UK: Cambridge University Press), 555–586. doi: 10.1017/CBO9780511691904.017

Berliner, D. C. (2002). Comment: Educational research: the hardest science of all. *Educ. Res.* 31, 18–20. doi: 10.3102/0013189X031008018

Besa, K.-S., Gesang, J., and Hinzke, J.-H. (2020). Zum Verhältnis von Forschungskompetenz und Unterrichtsplanung (On the relationship between research literacy and instructional design). *J. Allgem. Didak.* 10, 91–107.

Black, P., and Wiliam, D. (2003). 'In praise of educational research': formative assessment. *Brit. Educ. Res. J.* 29, 623–637. doi: 10.1080/0141192032000133721

Blöbaum, B. (2016). "Key factors in the process of trust. On the analysis of trust under digital conditions," in *Trust and Communication in a Digitized World*, ed B. Blöbaum (Cham: Springer International Publishing), 3–25. doi: 10.1007/978-3-319-28059-2_1

Bråten, I., and Ferguson, L. E. (2015). Beliefs about sources of knowledge predict motivation for learning in teacher education. *Teach. Teach. Educ.* 50, 13–23. doi: 10.1016/j.tate.2015.04.003

Bromme, R., and Goldman, S. R. (2014). The public's bounded understanding of science. *Educ. Psychol.* 49, 59–69. doi: 10.1080/00461520.2014.921572

Bromme, R., Kienhues, D., and Porsch, T. (2010). "Who knows what and who can we believe? Epistemological beliefs are beliefs about knowledge (mostly) to be attained from others," in *Personal Epistemology in the Classroom*, eds L. D. Bendixen and F. C. Feucht (Cambridge: Cambridge University Press), 163–194. doi: 10.1017/CBO9780511691904.006

Bromme, R., Mede, N. G., Thomm, E., Kremer, B., and Ziegler, R. (2022). An anchor in troubled times: trust in science before and within the COVID-19 pandemic. *PLoS ONE* 17, e0262823. doi: 10.1371/journal.pone.0262823

Bromme, R., Thomm, E., and Wolf, V. (2015). From understanding to deference: laypersons' and medical students' views on conflicts within medicine. *Int. J. Sci. Educ. B* 5, 68–91. doi: 10.1080/21548455.2013.849017

Brown, C. (2020). *The Networked School Leader: How to Improve Teaching and Student Outcomes Using Learning Networks*. Bingley, UK: Emerald Publishing. doi: 10.1108/9781838677190

Brown, C., MacGregor, S., Flood, J., and Malin, J. (2022). Facilitating research-informed educational practice for inclusion. Survey findings from 147 teachers and school leaders in England. *Front. Educ.* 7, 890832. doi: 10.3389/feduc.2022.890832

Brown, C., Schildkamp, K., and Hubers, M. D. (2017). Combining the best of two worlds: a conceptual proposal for evidence-informed school improvement. *Educ. Res.* 59, 154–172. doi: 10.1080/00131881.2017.1304327

Buehl, M. M., and Alexander, P. A. (2001). Beliefs about academic knowledge. *Educ. Psychol. Rev.* 13, 385–418. doi: 10.1023/A:1011917914756

Buehl, M. M., and Fives, H. (2009). Exploring teachers' beliefs about teaching knowledge: Where does it come from? Does it change? *J. Exp. Educ.* 77, 367–407. doi: 10.3200/JEXE.77.4.367-408

Bürkner, P.-C. (2017). Brms: an R package for Bayesian multilevel models using Stan. *J. Stat. Softw.* 80, 1–28. doi: 10.18637/jss.v080.i01

Butzer, B. (2020). Bias in the evaluation of psychology studies: a comparison of parapsychology versus neuroscience. *EXPLORE* 16, 382–391. doi: 10.1016/j.explore.2019.12.010

Cain, T. (2015). Teachers' engagement with published research: addressing the knowledge problem. *Curric. J.* 26, 488–509. doi: 10.1080/09585176.2015.1020820

Cochran-Smith, M., Ell, F., Ludlow, L., Grudnoff, L., and Aitken, G. (2014). The challenge and promise of complexity theory for teacher education research. *Teach. Coll. Rec.* 116, 1–38. doi: 10.1177/016146811411600407

Cramer, C. (2013). Beurteilung des bildungswissenschaftlichen Studiums durch Lehramtsstudierende in der ersten Ausbildungsphase im Längsschnitt (The assessment of educational scientific study courses as given by students in their first phase of teacher training in longitudinal section). *Z. Pädagog.* 59, 66–82. doi: 10.25656/01:11927

Deng, L., and Yuen, A. H. K. (2011). Towards a framework for educational affordances of blogs. *Comput. Educ.* 56, 441–451. doi: 10.1016/j.compedu.2010.09.005

Dietz, G., and Den Hartog, D. N. (2006). Measuring trust inside organisations. *Pers. Rev.* 35, 557–588. doi: 10.1108/00483480610682299

Doabler, C. T., Nelson, N. J., Kosty, D. B., Fien, H., Baker, S. K., Smolkowski, K., et al. (2014). Examining teachers' use of evidence-based practices during core mathematics instruction. *Assess. Effect. Intervent.* 39, 99–111. doi: 10.1177/1534508413511848

Druckman, J. N., and McGrath, M. C. (2019). The evidence for motivated reasoning in climate change preference formation. *Nat. Clim. Change* 9, 111–119. doi: 10.1038/s41558-018-0360-1

European Commission (2007). *Towards More Knowledge-Based Policy and Practice in Education and Training.* Available online at: https://op.europa.eu/en/publication-detail/-/publication/962e3b89-c546-4680-ac84-777f8f10c590 (accessed November 11, 2022).

Fives, H., and Buehl, M. M. (2010). "Teachers' articulation of beliefs about teaching knowledge: Conceptualizing a belief framework," in *Personal Epistemology in the Classroom*, eds L. D. Bendixen and F. C. Feucht (Cambridge, UK: Cambridge University Press), 470–515. doi: 10.1017/CBO9780511691904.015

Fives, H., and Buehl, M. M. (2012). "Spring cleaning for the "messy" construct of teachers' beliefs: What are they? Which have been examined? What can they tell us?," in *APA Educational Psychology Handbook, Vol 2: Individual Differences and Cultural and Contextual Factors*, eds K. R. Harris, S. Graham, T. Urdan, S. Graham, J. M. Royer, and M. Zeidner (Washington, DC: American Psychological Association), 471–499. doi: 10.1037/13274-019

Gelman, A., Goodrich, B., Gabry, J., and Vehtari, A. (2019). R-squared for Bayesian regression models. *Amer. Stat.* 73, 307–309. doi: 10.1080/00031305.2018.1549100

Gelman, A., and Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*, Vol. 1. New York, NY: Cambridge University Press. doi: 10.1017/CBO9780511790942

Gitlin, A., Barlow, L., Burbank, M. D., Kauchak, D., and Stevens, T. (1999). Pre-service teachers' thinking on research: implications for inquiry oriented teacher education. *Teach. Teach. Educ.* 15, 753–769. doi: 10.1016/S0742-051X(99)00015-3

Gough, D. (2021). Appraising evidence claims. *Rev. Res. Educ.* 45, 1–26. doi: 10.3102/0091732X20985072

Gronau, Q. F., Sarafoglou, A., Matzke, D., Ly, A., Boehm, U., Marsman, M., et al. (2017). A tutorial on bridge sampling. *J. Math. Psychol.* 81, 80–97. doi: 10.1016/j.jmp.2017.09.005

Groß Ophoff, J., and Cramer, C. (2022). "The engagement of teachers and school leaders with data, evidence, and research in Germany," in *The Emerald International Handbook of Evidence-Informed Practice in Education*, eds C. Brown and J. Malin (Bingley, UK: Emerald Publishing Limited), 175–195. doi: 10.1108/978-1-80043-141-620221026

Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., and Merrill, L. (2009). Feeling validated versus being correct: a meta-analysis of selective exposure to information. *Psychol. Bull.* 135, 555–588. doi: 10.1037/a0015701

Hendriks, F., Kienhues, D., and Bromme, R. (2015). Measuring laypeople's trust in experts in a digital age: the Muenster Epistemic Trustworthiness Inventory (METI). *PLoS ONE* 10, e0139309. doi: 10.1371/journal.pone.0139309

Hendriks, F., Kienhues, D., and Bromme, R. (2016). "Trust in science and the science of trust," in *Trust and Communication in a Digitized World*, ed B. Blöbaum (Cham: Springer International Publishing), 143–159. doi: 10.1007/978-3-319-28059-2_8

Hendriks, F., Seifried, E., and Menz, C. (2021). Unraveling the "smart but evil" stereotype: pre-service teachers' evaluations of educational psychology researchers versus teachers as sources of information. *Z. Pädagog. Psychol.* 35, 157–171. doi: 10.1024/1010-0652/a000300

Hinzke, J.-H., Gesang, J., and Besa, K.-S. (2020). Zur Erschließung der Nutzung von Forschungsergebnissen durch Lehrpersonen. Forschungsrelevanz zwischen Theorie und Praxis (Exploration of the use of research results by teachers. Relevance of research between theory and practice). *Z. Erziehungswiss.* 23, 1303–1323. doi: 10.1007/s11618-020-00982-6

Hofer, B. K., and Pintrich, P. R. (1997). The development of epistemological theories: beliefs about knowledge and knowing and their relation to learning. *Rev. Educ. Res.* 67, 88–140. doi: 10.3102/00346543067001088

Jonas, E., Schulz-Hardt, S., Frey, D., and Thelen, N. (2001). Confirmation bias in sequential information search after preliminary decisions: an expansion of dissonance theoretical research on selective exposure to information. *J. Pers. Soc. Psychol.* 80, 557–571. doi: 10.1037/0022-3514.80.4.557

Katz, S., and Dack, L. A. (2014). Towards a culture of inquiry for data use in schools: breaking down professional learning barriers through intentional interruption. *Stud. Educ. Eval.* 42, 35–40. doi: 10.1016/j.stueduc.2013.10.006

Kiemer, K., and Kollar, I. (2021). Source selection and source use as a basis for evidence-informed teaching. *Z. Pädagog. Psychol.* 35, 127–141. doi: 10.1024/1010-0652/a000302

Krammer, G., Vogel, S. E., and Grabner, R. H. (2021). Believing in neuromyths makes neither a bad nor good student-teacher: the relationship between neuromyths and academic achievement in teacher education. *Mind Brain Educ.* 15, 54–60. doi: 10.1111/mbe.12266

Krammer, G., Vogel, S. E., Yardimci, T., and Grabner, R. H. (2019). Neuromythen sind zu Beginn des Lehramtsstudiums prävalent und unabhängig vom Wissen über das menschliche Gehirn. *Z. Bildungsfors.* 9, 221–246. doi: 10.1007/s35834-019-00238-2

Kultusministerkonferenz (2014). *Standards für die Lehrerbildung: Bildungswissenschaften.* Available online at: https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung-Bildungswissenschaften.pdf (accessed November 11, 2022).

Landrum, T. J., Cook, B. G., Tankersley, M., and Fitzgerald, S. (2002). Teacher perceptions of the trustworthiness, usability, and accessibility of information from different sources. *Remed. Spec. Educ.* 23, 42–48. doi: 10.1177/074193250202300106

Lavrakas, P. (2008). *Respondent Fatigue.* Thousand Oaks, CA: Sage Publications. doi: 10.4135/9781412963947.n480

Lord, C. G., Lepper, M. R., and Preston, E. (1984). Considering the opposite: a corrective strategy for social judgment. *J. Pers. Soc. Psychol.* 47, 1231–1243. doi: 10.1037/0022-3514.47.6.1231

Lord, C. G., Ross, L., and Lepper, M. R. (1979). Biased assimilation and attitude polarization: the effects of prior theories on subsequently considered evidence. *J. Pers. Soc. Psychol.* 37, 2098–2109. doi: 10.1037/0022-3514.37.11.2098

Lüdtke, O., Robitzsch, A., and Grund, S. (2017). Multiple imputation of missing data in multilevel designs: a comparison of different strategies. *Psychol. Methods* 22, 141–165. doi: 10.1037/met0000096

Lüdtke, O., Robitzsch, A., Trautwein, U., and Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung. Probleme und Lösungen (Handling of missing data in psychological research. Problems and solutions). *Psychol. Rund.* 58, 103–117. doi: 10.1026/0033-3042.58.2.103

Makel, M. C., and Plucker, J. A. (2014). Facts are more important than novelty: replication in the education sciences. *Educ. Res.* 43, 304–316. doi: 10.3102/0013189X14545513

Mansournia, M. A., and Altman, D. G. (2016). Inverse probability weighting. *BMJ* 352, i189. doi: 10.1136/bmj.i189

Masnick, A. M., and Zimmerman, C. (2009). Evaluating scientific research in the context of prior belief: hindsight bias or confirmation bias? *J. Psychol. Sci. Technol.* 2, 29–36. doi: 10.1891/1939-7054.2.1.29

Mayer, R. C., Davis, J. H., and Schoorman, F. D. (1995). An integrative model of organizational trust. *Acad. Manage. Rev.* 20, 709–734. doi: 10.5465/amr.1995.9508080335

Menz, C., Spinath, B., and Seifried, E. (2021). Where do pre-service teachers' educational psychological misconceptions come from? *Z. Pädag. Psychol.* 35, 143–156. doi: 10.1024/1010-0652/a000299

Merk, S., and Rosman, T. (2019). Smart but evil? Student-Teachers' perception of educational researchers' epistemic trustworthiness. *AERA Open* 5, 233285841986815. doi: 10.1177/2332858419868158

Merk, S., Rosman, T., Muis, K. R., Kelava, A., and Bohl, T. (2018). Topic specific epistemic beliefs: extending the theory of integrated domains in personal epistemology. *Learn. Instruct.* 56, 84–97. doi: 10.1016/j.learninstruc.2018.04.008

Muis, K. R. (2004). Personal epistemology and mathematics: a critical review and synthesis of research. *Rev. Educ. Res.* 74, 317–377. doi: 10.3102/00346543074003317

Muis, K. R., Bendixen, L. D., and Haerle, F. C. (2006). Domain-generality and domain-specificity in personal epistemology research: philosophical and empirical reflections in the development of a theoretical framework. *Educ. Psychol. Rev.* 18, 3–54. doi: 10.1007/s10648-006-9003-6

Muis, K. R., Trevors, G., and Chevrier, M. (2016). "Epistemic climate for epistemic change," in *Handbook of Epistemic Cognition*, eds J. A. Greene, W. A. Sandoval, and I. Bråten (Abingdon: Routledge), 213–225.

Mullen, B., Brown, R., and Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: an integration. *Eur. J. Soc. Psychol.* 22, 103–122. doi: 10.1002/ejsp.2420220202

Nickerson, R. S. (1998). Confirmation bias: a ubiquitous phenomenon in many guises. *Rev. Gen. Psychol.* 2, 175–220. doi: 10.1037/1089-2680.2.2.175

Ostinelli, G. (2009). Teacher education in Italy, Germany, England, Sweden and Finland. *Eur. J. Educ.* 44, 291–308. doi: 10.1111/j.1465-3435.2009.01383.x

Oswald, M. E., and Grosjean, S. (2004). "Confirmation bias," in *Cognitive Illusions. A Handbook on Fallacies and Biases in Thinking, Judgement and Memory*, ed R. F. Pohl (Hove, NY: Psychology Press), 79–96.

Parr, J. M., and Timperley, H. S. (2008). Teachers, schools and using evidence: considerations of preparedness. *Assess. Educ. Princip. Pol. Pract.* 15, 57–71. doi: 10.1080/09695940701876151

Pashler, H., McDaniel, M., Rohrer, D., and Bjork, R. (2008). Learning styles: concepts and evidence. *Psychol. Sci. Publ. Int.* 9, 105–119. doi: 10.1111/j.1539-6053.2009.01038.x

Perry, W. G. (1999). *Forms of Intellectual and Ethical Development in the College Years: A Scheme. Jossey-Bass Higher and Adult Education Series*. Hoboken, NJ: Jossey-Bass.

Ray, B. B., and Hocutt, M. M. (2006). Teacher-created, teacher-centered weblogs: perceptions and practices. *J. Comput. Teach. Educ.* 23, 11–18. doi: 10.1080/10402454.2006.10784555

Renkl, A. (2022). Meta-analyses as a privileged information source for informing teachers' practice? A plea for theories as primus inter pares. *Z. Pädagog. Psychol.* 36, 217–231. doi: 10.1024/1010-0652/a000345

Rosman, T., Kerwer, M., Chasiotis, A., and Wedderhoff, O. (2021). Person- and situation-specific factors in discounting science via scientific impotence excuses. *Eur. J. Psychol.* 17, 288–305. doi: 10.5964/ejop.3735

Rosman, T., and Merk, S. (2021). Teacher's reasons for trust and distrust in scientific evidence: Reflecting a "smart but evil" stereotype? *AERA Open.* 7:23328584211028599. doi: 10.1177/23328584211028599

Rubin, D. B. (1976). Inference and missing data. *Biometrika* 63, 581–592. doi: 10.1093/biomet/63.3.581

Schmidt, K., Rosman, T., Cramer, C., Besa, K.-S., and Merk, S. (2022). *Teachers Trust Educational Science - Especially if it Confirms their Beliefs*. OSF. Available online at: https://osf.io/jm4tx/ (accessed November 11, 2022).

Shank, G., and Brown, L. (2007). *Exploring Educational Research Literacy*. New York, NY: Routledge.

Stan Development Team (2017). *Stan Modeling Language: User's Guide and Reference Manual*.

Stark, R. (2017). Probleme evidenzbasierter bzw.-orientierter pädagogischer Praxis [Problems of evidence-based or rather evidence-oriented educational practice]. *Z. Pädagog. Psychol.* 31, 99–110. doi: 10.1024/1010-0652/a000201

Statistisches Bundesamt (Destatis) (2019). Bildung, Forschung und Kultur (Education, research and culture). *Genesis*. (Online).

Stroud, N. J. (2017). "Understanding and overcoming selective exposure and judgment when communicating about science," in *The Oxford Handbook of the Science of Science Communication*, Vol. 1, eds K. H. Jamieson, D. M. Kahan and D. A. Scheufele (Oxford: Oxford University Press), 377–388. doi: 10.1093/oxfordhb/9780190497620.013.41

Taber, C. S., and Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *Amer. J. Polit. Sci.* 50, 755–769. doi: 10.1111/j.1540-5907.2006.00214.x

Thiem, J., Preetz, R., and Haberstroh, S. (2020). 'Warum soll ich forschen?' - Wirkungen forschenden Lernens bei Lehramtsstudierenden ('Why should I do research?': effects of research-based learning on student teachers). *Z. Hochschulentw.* 15, 187–207. doi: 10.3217/zfhe-15-02/10

Thomm, E., and Bromme, R. (2012). "It should at least seem scientific!" Textual features of "scientificness" and their impact on lay assessments of online information. *Sci. Educ.* 96, 187–211. doi: 10.1002/sce.20480

Thomm, E., Gold, B., Betsch, T., and Bauer, J. (2021a). When preservice teachers' prior beliefs contradict evidence from educational research. *Brit. J. Educ. Psychol.* 91, e12407. doi: 10.1111/bjep.12407

Thomm, E., Sälzer, C., Prenzel, M., and Bauer, J. (2021b). Predictors of teachers' appreciation of evidence-based practice and educational research findings. *Z. Pädagog. Psychol.* 35, 173–184. doi: 10.1024/1010-0652/a000301

van Buuren, S. (2018). *Flexible Imputation of Missing Data*. Interdisciplinary Statistics Series. Boca Raton, FL: Chapman and Hall/CRC Press, Taylor and Francis Group.

Van Lommel, K., Van Gasse, R., Van Hoof, J., and Van Petegem, P. (2017). Teachers' decision-making: data based or intuition driven? *Int. J. Educ. Re.*, 83, 75–83. doi: 10.1016/j.ijer.2017.02.013

van Schaik, P., Volman, M., Admiraal, W., and Schenke, W. (2018). Barriers and conditions for teachers' utilisation of academic knowledge. *Int. J. Educ. Res.* 90, 50–63. doi: 10.1016/j.ijer.2018.05.003

Weißkopf, M., Ziegler, R., and Kremer, B. (2019). *Science Barometer 2018*. Cologne: GESIS Data Archive.

Williams, D., and Coles, L. (2007). Teachers' approaches to finding and using research evidence: an information literacy perspective. *Educ. Res.* 49, 185–206. doi: 10.1080/00131880701369719

Wolter, K., Chowdhury, S., and Kelly, J. (2009). "Chapter 7 - Design, conduct, and analysis of random-digit dialing surveys," in *Handbook of Statistics, Vol. 29*, ed C. R. Rao (Amsterdam: Elsevier), 125–154. doi: 10.1016/S0169-7161(08)00007-2

Zeuch, N., and Souvignier, E. (2016). "Wissenschaftliches Denken bei Lehramts- und Psychologiestudierenden (Scientific thinking in teacher and psychology students)," in *Psychologiedidaktik und Evaluation XI (Psychology Didactics and Evaluation XI)*, eds M. Krämer, S. Preiser, and K. Brusdeylins (Aachen: Shaker Verlag), 175–183.

Zhou, X., and Reiter, J. P. (2010). A note on Bayesian inference after multiple imputation. *Amer. Statist.* 64, 159–163. doi: 10.1198/tast.2010.09109