HISTORICAL ENCOUNTERS
A journal of historical consciousness, historical cultures, and history education

ISSN 2203-7543 | Double Blind Peer-Reviewed | Open Access | www.hej-hermes.net

Epistemic beliefs of Norwegian history student teachers: Testing and assessing two measurement instruments

David-Alexandre Wagner
University of Stavanger, Norway

Ulrich Dettweiler
University of Stavanger, Norway

ABSTRACT

Since teachers' and students' epistemic beliefs about history are believed to significantly impact teachers' practices, students' performance and the ability to think historically of both, investigating such beliefs is important. Following the seminal works of Maggioni and colleagues (Maggioni, 2010; Maggioni et al., 2004; Maggioni et al., 2009), a number of studies have adapted versions of her Beliefs About Learning and Teaching History Questionnaire (BLTHQ) and Beliefs about History Questionnaire (BHQ) to quantitatively evaluate epistemic beliefs in different national contexts (Mierwald et al., 2016; Miguel-Revilla et al., 2017; Namamba & Rao, 2016; Nitsche, 2019; Stoel, Logtenberg, et al., 2017). However, the validity and reliability of these instruments have mostly been found to be problematic (Mierwald & Junius, 2022; Stoel et al., 2022). We have tested two different questionnaires – Maggioni's BLTHQ and our own adaptation of the BHQ – in separate surveys, on respectively 176 and 324 Norwegian history student teachers. For both questionnaires, our exploratory analysis showed a three-factor solution, supporting King and Kitchener's (1994) model, and undermining the two-factor solutions found in many former studies. However, despite both questionnaires showing acceptable consistency and fitness, these levels were not optimal, and confirmed the problem of epistemic inconsistency (or 'wobbling') revealed by previous studies. These results encourage further qualitative studies to better understand the problem of wobbling, to design better questionnaires to be tested in the future.

KEYWORDS

Epistemic beliefs, Epistemic wobbling, Norwegian history student teachers, Measurement instruments

CITATION

Wagner, D.-A., & Dettweiler, U. (2024). Epistemic beliefs of Norwegian history student teachers: Testing and assessing two measurement instruments, *Historical Encounters, 11*(2), 98-113. https://doi.org/10.52289/hej11.207

Introduction

Since the 1970s, epistemic beliefs have been widely acknowledged as having significant influence upon, and partially predicting, students' and teachers' metacognitive processes to learn and teach and their levels of critical thinking in different domains of knowledge (Buehl & Alexander, 2001; Hofer & Bendixen, 2012; King & Kitchener, 2002; Kuhn, 1999; Stoel et al., 2022; VanSledright & Limón, 2006). Consequently, in History and Social Studies education, a range of different models have been conceptualized and tested to understand and evaluate people's epistemic beliefs (Maggioni et al., 2009; Miguel-Revilla et al., 2020; Stoel et al., 2022). The Beliefs about Learning and Teaching of History Questionnaire (BLTHQ) developed by Liliana Maggioni and colleagues (Maggioni et al., 2004; Maggioni et al., 2009) and its later version, the Beliefs about History Questionnaire (BHQ) (2010) have gained particular interest, as they reveal latent structures of directly obtained beliefs and offer a quantitative approach that saves considerable time and energy, compared with some more nuanced yet more complex qualitative surveys. However, both questionnaires have proven to meet several challenges, which are discussed in detail below (Maggioni, 2010; Maggioni et al., 2004; Maggioni et al., 2009; Mierwald et al., 2017; Stoel, van Drie, & van Boxtel, 2017).

Following Maggioni's work, our goal has been to test two versions of the BLTHQ in a Norwegian context. To do so, we have translated the original BLTHQ into Norwegian, and developed a modified version (hereby called OHF) derived from the BHQ and other similar questionnaires. We have given both questionnaires to Norwegian history student teachers and compared their respective factorial structures, and tested their validity.

Theoretical framework

The role of epistemic beliefs about history in education

In the wake of the cognitive revolution started in the 1950s' (Royer, 2006) and the Schools Council History Project in the UK in the 1970s' (Shemilt, 1980), we can see a paradigmatic shift in the way history education is conceptualized. Namely, a shift from history studies focusing on the memorization of a set of facts and national narratives, to the recognition of history as encapsulating a diversity of topics, perspectives, methodologies, skills and abilities (Lévesque, 2008; Mathis & Parkes, 2020). In particular, the skills and competences associated with "doing history" have provoked questions about students' and teachers' epistemic ideas about history and historical knowledge (Mathis & Parkes, 2020). Furthermore, the level of students' epistemic beliefs has been closely linked to the achievement of greater disciplinary motivation, academic performance, and critical thinking (Buehl & Alexander, 2001; Lee & Shemilt, 2003). In turn, these insights have been deemed crucial for the proper participation of future citizens in pluralistic and democratic societies where both the role of social media and the flow of information, are increasing (Barton & Levstik, 2009; Journell, 2017).

Regarding teachers, it is upheld that epistemic beliefs influence their teaching practices and their ability to change them durably (Brownlee et al., 2017; Buehl & Fives, 2016; Hofer & Bendixen, 2012, pp. 239-241).

Defining and assessing epistemic beliefs about history

The theoretical conceptualization of epistemic beliefs in history education has been mainly informed and influenced by the models developed by King and Kitchener (King & Kitchener, 1994, 2002), Kuhn and Weinstock (Kuhn, 1999; Kuhn & Weinstock, 2002), and Lee and Shemilt (2003). Through qualitative investigation of tasks and in-depth interviews, these models have outlined different stages of development, where the highest level is the ideal to attain. Maggioni and colleagues have aimed to combine and synthesize these models to design a questionnaire – the BLTHQ – that could be used to quantitatively determine people's epistemic beliefs on a three-level scale (Maggioni et al., 2004; Maggioni et al., 2009). At the first level, the copier stance is equivalent to King & Kitchener's pre-reflective period and Kuhn and Weinstock's realist and absolutist positions (schematically, historical knowledge is considered as objective and mirroring "what happened"). At the second level, the borrower stance corresponds to King and Kitchener's quasi-reflective period and Kuhn and Weinstock's multiplist positions (historical knowledge is considered as subjective and relative). At the highest level of progression, the reflective stance parallels King and Kitchener's reflective period and Kuhn & Weinstock's evaluativist positions (historical knowledge is the result of constructed interpretations and narratives based on debated disciplinary criteria of scientificity). The original BLTHQ was replaced in 2010 by an alternative, the BHQ (Maggioni, 2010).

The design of the BLTHQ and the BHQ were well-grounded and led to German and Spanish translations (Mierwald et al., 2017; Miguel-Revilla & Fernández, 2017). However, issues arose regarding validity and the interpretation of its answers. First, the factor analysis of the original BLTHQ isolated only two factors: one including the objectivist and subjectivist items, the other encompassing the criterialist items – instead of the three factors expected in the light of the original theoretical model and King & Kitchener's Reflective Judgment Model (RJM). The same issue arose in the first version of the German BHQ (22 items – a study called ARGUMENT), tested on 124 upper secondary school pupils (Mierwald et al., 2017). Moreover, the loadings of some items, particularly the objectivist and the criterialist ones, were problematic in the German, Dutch and Spanish translations of the BHQ (Mierwald & Junius, 2022; Miguel-Revilla et al., 2020; Stoel, Logtenberg, et al., 2017). Consequently, the theoretical background of the subjectivist items has been criticized as ambiguous and potentially wobbling between naïve and more nuanced statements (Mierwald & Junius, 2022; Stoel, Logtenberg, et al., 2017).

Stoel, Logtenberg, et al. (2017) tested an alternative questionnaire of 26 items, partly based on a Dutch translation of the BHQ, on 922 upper secondary school students, aiming to highlight a two-factor solution between naïve beliefs (15 items – beliefs viewing the past as fixed or as the result of opinions) and nuanced epistemic beliefs (11 items – seeing history as multiple interpretations of the past and the result of disciplinary criteria). As a result, five factors emerged from their exploratory factor analysis (EFA), and although some factors could be categorized as belonging either to naïve or nuanced epistemic beliefs, the original aim of obtaining a clear distinction between two factors proved difficult to achieve.

Moreover, an issue of epistemic inconsistency has often been identified, particularly for beginners (Maggioni, 2010), but also among experts (Stoel, Logtenberg, et al., 2017). This shows that participants may hold views reflecting different stances that appear to be contradictory, for example agreeing (or disagreeing) with statements belonging either to the pre-reflective/objectivist stance, the quasi-reflective/relativist/subjectivist stance or the reflective/criterialist/evaluativist stance.

Interestingly, Mierwald et al. (2017) tested a modified version of the BHQ (a study called SOSCIE) on 224 German student teachers. They obtained originally a six-factor structure, that was reduced to three factors, based on the visual inspection of the Scree plot. Those three factors explained 41% of the total variance and most items intended to each stance loaded on the same factor – in accordance with Maggioni's (2010) three epistemic stances – and with good internal consistency (Mierwald, 2020; Mierwald & Junius, 2022; Mierwald et al., 2017). However, the SOSCIE-study had reduced the original Likert-scale from six to four alternatives and forced the

confirmatory factor analysis (CFA) with three factors. Reducing the complexity in representation of the respondents' beliefs may have overestimated the effects and makes comparisons difficult.

Some years later, Mierwald and Junius (2022) performed a think-aloud questionnaire and interviews with four German students. Their study showed that, although most items in the questionnaire were clear and easy to understand, a few statements from the criterialist and objectivist stances were still problematic. Together with previous studies, this supports reducing the complexity and the epistemic ambiguity of the questionnaire, by rewording the statements more adequately, fastening them closer to the stance they mean to represent, and discard references to the school context (Barzilai & Weinstock, 2015; Mierwald & Junius, 2022; Muis et al., 2014).

In addition, several studies have indicated that the national and cultural context of each country may influence the way people understood and answered questions related to their epistemic conceptions about history (Sakki & Pirttilä-Backman, 2019; Stoel et al., 2022; Wansink et al., 2016).

Therefore, there are many interesting questions to consider when adapting and testing the BLTHQ and a modified version of the BHQ:

- Does a replication of the BLTHQ among Norwegian history student teachers show results comparable to those from the original study (Maggioni et al., 2004; Maggioni et al., 2009) and the ARGUMENT-study (Mierwald et al., 2017)?
- Does a test of the validity of the BLTHQ and of a modified version of the BHQ in the Norwegian context show significantly different results – for example, a better fit to a three-factor distribution and a solid consistency of the different factors?

Method

We tested two different questionnaires in separate surveys. The first survey used a translation of the BLTHQ (Maggioni et al., 2004) in Norwegian. The second survey used a questionnaire we have developed (called *Oppfatninger om historiefaget* [Beliefs About History as a Discipline], OHF – see Table 1), based on a mix of Maggioni's BLTHQ (2004) and BHQ (2010), and King and Kitchener's principles for their RJM (1994). Following Stoel et al. (2022), we considered the stages in the different developmental models to be roughly equivalent. Table 2 shows a comparison between the BLTHQ, the BHQ and our OHF. Together with introducing some nuance in the wording of certain statements, we made changes in the hope of better singling out the different groups.

Firstly, we chose to have an equal number of items in each category. The BLTHQ and the BHQ have an unbalanced number of statements for each stance, and we wanted to avoid any imbalances which could potentially impact the results. The BLTHQ included a total of 21 items, nine of them belonging to the objectivist stance; eight to the subjectivist stance; and four to the criterialist stance. The BHQ was originally composed of 22 items; five items in the copier stance; nine items in the borrower stance; and eight items in the criterialist stance. We opted for a total of 18 items, assigning six items to each stance, making the questionnaire more compact and balanced.

Secondly, some statements in the original questionnaires seemed redundant or to be overlapping (Table 2). For example, "Students who are good at memorization learn history quickly" (BLTHQ – item 1) and "To learn history means mainly to study many facts about the past and commit them to memory" (BLTHQ – item 6); or "History is simply a matter of interpretation" (BHQ – item 2) and "Students need to be aware that history is essentially a matter of interpretation" (BHQ – item 17); or "A historical account is the product of a disciplined method of inquiry" (BHQ – item 3) and "History is a critical inquiry about the past" (BHQ – item 11). We tried to avoid questions or statements that were too similar.

Thirdly, we kept the item S4 "History should be taught like a story: Some things are true, but some others are just a matter of personal opinion" (Table 2), but moved it from the subjectivist

stance (in the BLTHQ) to the objectivist stance in our OHF-model, considering that this item (P5) may better correspond to the description in Stage 3 of the RJM – stating that knowledge is assumed to be certain, but in case of uncertainty, personal opinions will prevail (King & Kitchener, 1994, pp. 56-57).

Table 1

Items defined in our OHF-questionnaire, grouped by stance

Code	Item and item number
	Objectivist/Pre-reflective stance
P1	1. History is not only about learning and memorizing facts, but it is the most important part.
P2	2. The facts speak often for themselves in history and do not need to be discussed or debated.
P3	3. It is fully possible to be objective in History, if one examines things with an open mind.
P4	6. Good general reading and comprehension skills are usually enough to learn history.
P5	14. History should be taught as a narrative: Certain things are true, and others are only a matter of personal opinions.
P6	18. When in doubt between to contradicting interpretations, one should choose the interpretation coming from the person one trusts the most.
	Subjectivist/Quasi-reflective stance
Q1	4. What we know in history is relative: Two contradicting interpretations can both be true, depending on the perspective.
Q2	8. The choice of arguments and evidence in history is subjective and individual.
Q3	9. When reading an historical account, it is more important to focus on the author's perspective than on how he/she supports their reasoning.
Q4	10. It is actually impossible to be sure of anything in history: one can question most of it.
Q5	11. To teach/learn that one particular interpretation is better than another is in fact unfortunate in history.
Q6	13. In history books, the accounts are usually largely depending on the historian's own perspective.
	Criterialist/Reflective stance
R1	5. The hope of being objective in history must be abandoned; one can only be aware of one's own choices of methods and perspectives.
R2	7. It is fundamental that students learn to assess methods used in history.
R3	12. Comparing sources and understanding author perspective is essential in history.
R4	15. In history, it is essential that students learn to deal with conflicting evidence/sources.
R5	16. It is fundamental that students are taught to support their reasoning with evidence and sources.
R6	17. History is about assessing the sources available to produce the most probable and trustworthy interpretation.
	Summarizing question
S	19. Which of the following alternatives is closest to your own conception of history as a discipline?
O	a) History is mainly about what is true and false; to find out what actually happened and why. One has to be as objective as possible, which is a difficult task, but one has to try.
Q	c) It is actually impossible to know anything for sure in history, because everything is subjective, relative and hinging on perspectives and points of view. One shall only understand the different perspectives, preferably without taking sides.
R	b) History is about understanding the past, based on different sources and remains available. It is about assessing perspectives and methods used to produce the most likely and trustworthy interpretation.

Table 2

Comparison of the three different models: BLTHQ, BHQ and OHF

BLTHQ (Maggioni et al. 2004, 2009)	BHQ (Maggioni 2010, Miguel-Revilla 2020, Mierwald et al. 2017)	OHF
Copier/Objectivist	**Copier/Objectivist**	**Objectivist/Pre-reflective stance**
O1 Students who are good at memorization learn history quickly.		**P1** History is not only about learning and memorizing facts, but it is the most important part.
O2 Corroborating evidence and identifying sources are important learning strategies in history, but only after mastering the basic facts		
O3 In history there is really nothing to understand; the facts speak for themselves.	16. The facts speak for themselves.	**P2** The facts speak often for themselves in history and do not need to be discussed or debated.
O4 Students who know their textbook well will be good at history.		
O5 To learn history means mainly to study many facts about the past and commit them to memory.		
O6 Teachers need to avoid giving students conflicting sources, since it makes historical investigation impossible.		
O7 In learning history, summarizing is more important than comparing.		
O8 Teachers should not question students' historical opinions, only check that they know the facts.	20. Teachers should not question students' historical opinions, only check that they know the facts.	
O9 Good general reading and comprehension skills are enough to learn history well.	9. Good general reading and comprehension skills are enough to learn history well.	**P4** Good general reading and comprehension skills are enough to learn history well.
	5. Disagreement about the same event in the past is always due to lack of evidence.	
		P3 It is fully possible to be objective in History, if one examines things with an open mind.
		P5 History should be taught like a story: Certain things are true, and others are only a matter of personal opinions.
		P6 When in doubt between two contradicting interpretations, one should choose the interpretation coming from the person one trusts the most.
	19. Even eyewitnesses do not always agree with each other, so there is no way to know what happened.	
	23. Differences in historical accounts result from absence or falsity of historical facts.*	
	25. History consists of the sum of collected historical facts.*	
Subjectivist/Quasi-reflective stance	**Subjectivist/Quasi-reflective stance**	**Subjectivist/Quasi-reflective stance**
S1 Students who read many history books learn that the past is what the historian makes it to be.	4. Students who read many history books learn that the past is what the historian makes it to be.	**Q6** In history books, the accounts are usually largely depending on the historian's own perspective.
S2 Good students know that history is basically a matter of opinion.	6. Good students know that history is basically a matter of opinion.	
S3 Students need to be aware that history is essentially a matter of interpretation.	17. Students need to be aware that history is essentially a matter of interpretation.	
S4 History should be taught like a story: Some things are true, but some others are just a matter of personal opinion.		
S5 In reading a history book, it is more important to pay attention to the perspective of the historian than to his or her reasoning on the evidence discussed.		**Q3** In reading a history book, it is more important to pay attention to the perspective of the historian than to his or her reasoning on the evidence discussed.
S6 Since there is no way to know what really happened in the past, students can believe whatever story they choose.	10. Since there is no way to know what really happened in the past, students can believe whatever story they choose.	
S7 Teaching that one historical interpretation is better than another is usually inappropriate.		**Q5** Teaching/Learning that one historical interpretation is better than another is usually inappropriate.
S8 Teachers need to make all historical interpretations available and let the students construct their own understanding of them.		
		Q1 What we know in history is relative: Two contradicting interpretations can both be true, depending on the perspective.
		Q2 The choice of arguments and evidence in history is subjective and individual.
	14. It is impossible to know anything for sure/with certainty about the past, since no one of us was there.	**Q4** It is actually impossible to be sure of anything in history: one can question most of it.
	2. History is simply a matter of interpretation	
	8. Historical claims cannot be justified/substantiated, since they are simply a matter of interpretation.	
	12. The past is what the historian makes it to be.	
	22. There is no evidence in history	
Criterialist stance	**Criterialist stance**	**Criterialist stance/Reflective stance**
C1 Knowledge of the historical method is fundamental for historians and students alike.	15. Knowledge of the historical method is fundamental for historians and students alike.	**R2** It is fundamental that students learn to assess the methods used in history.
C2 Comparing sources and looking for author subtext are essential components of the process of learning history.	13.Comparing sources and understanding author perspective are essential components of the process of learning history.	**R3** Comparing sources and understanding author perspective is essential in history.
C3 Students need to be taught to deal with conflicting evidence.	7. Students need to be taught to deal with conflicting evidence.	**R4** In history, it is essential that students learn to deal with conflicting evidence/sources.
C4 It is fundamental that students are taught to support their reasoning with evidence and ask that history textbook authors do so also.	1. It is fundamental that students are taught to support their reasoning with evidence.	**R5** It is fundamental that students are taught to support their reasoning with evidence and sources.
	21. History is the reasonable reconstruction of past occurrences based on the available evidence.	**R6** History is about assessing the sources available to produce the most probable and trustworthy interpretation.
		R1 The hope of being objective in history must be abandoned; one can only be aware of one's own choices of methods and perspectives.
	3. A historical account is the product of a disciplined method of inquiry.	
	11. History is a critical inquiry about the past.	
	18. Reasonable accounts can be constructed even in the presence of conflicting evidence.	
	24. Historians reconstruct the past based on regulated methods. *	

* Questions added to the original BHQ in Mierwald et al. (2017)

Moreover, the BLTHQ focused on creating statements affiliated to history learning and teaching, assuming that they were a good proxy for the epistemic beliefs of their main target: history teachers. Following recommendations from Mierwald and Junius (2022), we replaced recurrent references to the school system by more direct statements about conceptions about history as a discipline. For example, we added statements like "It is fully possible to be objective in History, if one examines things with an open mind", and "What we know in history is relative: Two contradicting interpretations can both be true, depending on the perspective", or "The hope of being objective in history must be abandoned; one can only be aware of one's own choices of methods and perspectives". We hoped that these additions, together with the inclusion of adverbs that introduce further nuance in the statements, would help the participants answer more easily and would make clearer divides between the three stance categories. In the same spirit, we added a last question that intended to explicitly summarize the meaning of each position, by asking each participant which stance was closest to their own overall epistemological position towards history (Table 1). This summarizing question is intended 1) to indicate clearly which stance each respondent thinks s/he belongs to, and 2) to display eventual disparities between the respondents' answers to the previous statements and their perception of their own overall epistemic position.

Finally, we kept the original six-point Likert scale system, as used by Maggioni (1 = Strongly disagree; 2 = Disagree; 3 = Somewhat disagree; 4 = Somewhat agree; 5 = Agree; 6 = Strongly agree). The order of the 18 different items was randomly set but was the same for all participants. The questionnaire was distributed online, using SurveyXact, through the University portal for History student teachers at different levels (from first year to master), from four different universities. We followed the formal ethical rules of confidentiality and personal data protection endorsed by the Norwegian Agency for Shared Services in Education and Research.

176 subjects (74 female and 102 male) filled in the BLTHQ – 215 were between 18 and 24 years old; their mean study experience was 2,3 years (median = 2,00; std. dev. = 1,60), and 155 were in their first year of History study. For our OHF-questionnaire, we gathered answers from 324 respondents (134 female and 190 male); 215 were between 18 and 24 years old; their mean study experience was 2,93 years (median = 3,00; std. dev. = 1,84); 200 were in their first year of study.

We have used JASP (2022, Version 0.16.2), a free software based on R, to extract our results and test the two different models through an EFA, and then a CFA. We have supposed that the different factors were not independent and used a Varimax rotation based on main components, excluding loadings lower than 0.400. Factor analysis is a statistical method used to measure whether a large number of variables (e.g. "items" in a questionnaire) can be reduced into fewer groups (e.g. here "objectivist", "subjectivist" or "criterialist" stances); the result for each item reflects how strongly it relates to a particular factor.

Findings and interpretation

The measure of the Kayser-Meyer-Olkin coefficient for both questionnaires turned out satisfactory to run a factor analysis (KMO for BLTHQ = .750; KMO for OHF = .726), which was also confirmed by the Bartlett's tests of sphericity – testing whether a matrix (of correlations) is significantly different from an identity matrix, it provides probability that the correlation matrix has significant correlations among at least some of the variables in a dataset, which is a prerequisite for factor analysis to work.

Table 3

Bartlett's test of sphericity

	X^2	df	p
BLTHQ	910.631	210.000	<.001
OHF	916.537	153.000	<.001

For both questionnaires, we managed to extract a three-factor solution through the EFA with a Varimax rotation (Tables 6 and 7). However, the three factors stood for a total of explanation of the variance of only 32.4% in the BLTHQ (Table 8), and for 27% in the OHF-questionnaire (Table 9). In both cases, furthermore, Factor 3 explained only 5,2% of the variance, which can be considered rather low. Moreover, although each factor clustered items belonging to the same stance, the hierarchy of the explaining factors in each model is not the same. In the OHF (Tables 5 and 7), Factor 1 gathers reflective/criterialist items (α =.64), Factor 2 cumulates quasi-reflective/subjectivist items (α =.53) and the less explaining Factor 3 associates pre-reflective/objectivist items (α =.57). In contrast, in the BLTHQ (Tables 4 and 6), Factor 1 gathers also criterialist items (α =.78), while Factor 2 gathers objectivist items (α =.68) and the less explaining Factor 3 cumulates the subjectivist items (α =.56). We used a Maximum Likelihood algorithm, and the measures of Unidimensional Reliability for the different scales show that, for the BLTHQ (Table 4), the reliability of the O-scale is acceptable, while the S-scale is weak and the C-scale is good. For the OHF, the reliability of the P-scale and the Q-scale is weak, while the R-scale is acceptable (Table 5). Besides, the p value of the Chi-squared test for both questionnaires was significant ($p < .001$), which indicates that both models do not fit optimally.

Moreover, although the Cronbach's α of the different factors in the EFA supports that the BLTHQ is more consistent than the OHF, the fit measures calculated in the CFA signal that the OHF fits better than the BLTHQ – the p-value of their Chi-square tests were respectively .032 versus .003; the value of their respective Comparative Fit Index (CFI) was .945 versus .923, while their respective Tucker-Lewis Index (TLI) showed values of .931 versus .905; the Root mean square error of approximation (RMSEA) of the OHF-questionnaire (.035) was lower than the RMSEA of the BLTHQ (.054). This apparent incongruity will be addressed in the discussion.

Table 4

Frequentist Unidimensional Scale Reliability for BLTHQ

	O-scale		S-scale		C-scale	
Estimate	McDonald's ω	Cronbach's α	McDonald's ω	Cronbach's α	McDonald's ω	Cronbach's α
Point estimate	0.685	0.676	0.562	0.557	0.781	0.778
95% CI lower bound	0.613	0.592	0.453	0.429	0.725	0.714
95% CI upper bound	0.757	0.745	0.671	0.660	0.837	0.829

Table 5

Frequentist Unidimensional Scale Reliability for OHF

	P-scale		Q-scale		R-scale	
Estimate	McDonald's ω	Cronbach's α	McDonald's ω	Cronbach's α	McDonald's ω	Cronbach's α
Point estimate	0.593	0.567	0.541	0.533	0.634	0.635
95% CI lower bound	0.521	0.486	0.460	0.448	0.568	0.565
95% CI upper bound	0.666	0.638	0.621	0.608	0.699	0.696

From Tables 6 and 7, we can observe that most items clustered as expected to their respective stance. However, some items did not: O6, O9, S8 and C1 were problematic regarding the BLTHQ (Table 6); and R1, P4, P5, P6, Q1 and Q6 were problematic concerning the OHF (Table 7). In addition, since the OHF-questionnaire and the BLTHQ have some statements in common, a comparison of their results is also interesting (Table 10). In both cases, there were three kinds of problem: 1) some items loaded negatively on a factor; 2) certain items' loadings were lower than 0.400; 3) certain similar items loaded very differently: higher than 0.400 in one questionnaire and lower than 0.400 in the other.

1) In the BLTHQ (Tables 6 and 10), O6 loaded negatively (-0.448) on factor 1 with the criterialist items, which means that O6 was a statement that criterialist informants disagreed with more than it was an item objectivist informants agreed with. Its negative formulation ("Teachers

need to avoid giving students conflicting sources, since it makes historical investigation impossible") may explain this result and should be revised so that it more clearly relates to the objectivist stance.

The same issue affected item R1 (Tables 7 and 10): R1 loaded negatively on factor 3 (which clustered objectivist items), which means that R1 was more a statement objectivist informants disagreed with than an item criterialist informants resolutely agreed with. This may be explained by the statement's formulation ("The hope of being objective in history must be abandoned; one can only be aware of one's own choices of methods and perspectives"): the first part has a clear (negative) connection to the objectivist stance that may have overshadowed the main intended criterialist nature of the statement.

Table 6

Factor Loadings – BLTHQ in Norwegian

	Factor 1	Factor 2	Factor 3	Uniqueness
C3	0.767			0.405
C4	0.710			0.496
C2	0.671			0.513
O6	-0.448			0.638
O5		0.726		0.423
O7		0.543		0.578
O3		0.526		0.678
O4		0.490		0.746
O1		0.470		0.746
O8		0.457		0.690
O2		0.413		0.790
S2			0.626	0.582
S1			0.602	0.626
S3			0.590	0.559
S7			0.453	0.754
S6			0.425	0.777
S5			0.424	0.782
S4			0.411	0.807
O9				0.886
S8				0.886
C1				0.839

Note. Applied rotation method is varimax.

Table 7

Factor Loadings – OHF-Questionnaire

	Factor 1	Factor 2	Factor 3	Uniqueness
R5	0.627			0.605
R4	0.591			0.590
R3	0.573			0.636
R6	0.494			0.735
R2	0.481			0.763
Q3		0.538		0.667
Q4		0.490		0.722
Q5		0.460		0.778
Q2		0.444		0.793
P3			0.667	0.552
R1			-0.559	0.602
P1			0.509	0.629
P2			0.448	0.671
P4				0.906
P5				0.899
P6				0.860
Q1				0.873
Q6				0.852

Note. Applied rotation method is varimax.

2) Items O9 and S8 were very unique, with results lower than 0.400 in the BLTHQ (Table 6); this was similar to items P4, P6 and Q1 in the OHF (Table 7). It means that they were not connected strongly enough to the stance they were meant to relate to. We can gather that their formulations were not contentious enough; these statements were easy to agree (or disagree) with, independently of the stance the respondents mainly belonged to.

Item O9 ("Good general reading and comprehension skills are enough to learn history well") in the BLTHQ was identical to item P4 in the OHF (Table 10). Their respective results were lower than 0.400, probably for the same reason: the statement was too widely formulated and not linked specifically enough to the objectivist stance. Respondents belonging to the other stances may also agree with this statement.

Likewise, item S8 ("Teachers need to make all historical interpretations available and let the students construct their own understanding of them") in the BLTHQ (Tables 7) was too widely formulated and linked specifically enough to the subjectivist stance.

Items P6 and Q1 (Table 7) have suffered of the same type of problem. P6 ("When in doubt between two contradicting interpretations, one should choose the interpretation coming from the person one trusts the most.") was meant to be understood as an objectivist statement. However,

the last part of the statement ("the person *one trusts the most*") is too ambiguous to be interpreted as clearly objectivist: the *trust* may be grounded on scientifically well-founded criteria and not only on the reputation of the person. Those belonging to the other stances may also agree with this statement. As for Q1 ("What we know in history is relative: Two contradicting interpretations can both be true, depending on the perspective."), the formulation is not ambiguous, but not exclusively subjectivist enough: both subjectivist and criterialist respondents may have agreed upon this statement.

3) Three similar sets of items had very different results (Table 10).

S1 ("Students who read many history books learn that the past is what the historian makes it to be") loaded significantly on factor 3, while Q6's ("In history books, the accounts are usually largely depending on the historian's own perspective.") loading is lower than .400. Table 7 shows that Q6's uniqueness is high.

The same issue appeared for C1 ("Knowledge of the historical method is fundamental for historians and students alike") and R2 ("It is fundamental that students learn to assess the methods used in history"). This may indicate a hermeneutical problem: the understanding of certain similar statements may differ substantially because of (small) wording differences, or because of inconsistencies in the participants' answers and epistemic conceptions. It may also be due to contextual reasons, such as the order of the items in the questionnaire.

As for S4 and P5 ("History should be taught like a story: Some things are true, but some others are just a matter of personal opinion"), it may be argued that their respective results are not that distant: S4's loading (0.411) is only just above 0.400 and the uniqueness of both items is high (Tables 6 and 7).

Table 8

Factor Characteristics – BLTHQ in Norweigan

	Unrotated solution			Rotated solution		
	SumSq. Loadings	Proportion var.	Cumulative	SumSq. Loadings	Proportion var.	Cumulative
Factor 1	3.468	0.165	0.165	2.397	0.114	0.114
Factor 2	2.244	0.107	0.272	2.352	0.112	0.226
Factor 3	1.086	0.052	0.324	2.048	0.098	0.324

Table 9

Factor Characteristics – OHF-Questionnaire

	Unrotated solution			Rotated solution		
	SumSq. Loadings	Proportion var.	Cumulative	SumSq. Loadings	Proportion var.	Cumulative
Factor 1	2.421	0.134	0.134	1.868	0.104	0.104
Factor 2	1.507	0.084	0.218	1.579	0.088	0.192
Factor 3	0.939	0.052	0.270	1.420	0.079	0.270

Table 10

Comparison of the loadings of the items in the two questionnaires

BLTHQ				OHF-questionnaire			
Factor 3	**Factor 2**	**Factor 1**			**Factor 1**	**Factor 2**	**Factor 3**
	0.470		O1	P1			0.509
	0.413		O2				
	0.526		O3	P2			0.448
	0.490		O4				
	0.726		O5				
		-0.448	O6				
	0.543		O7				
	0.457		O8				
			O9	P4			
				P3			0.667
				P6			
0.602			S1	Q6			
0.626			S2				
0.590			S3				
0.411			S4	P5			
0.424			S5	Q3		0.538	
0.425			S6				
0.453			S7	Q5		0.460	
			S8				
				Q1			
				Q2		0.444	
				Q4		0.490	
			C1	R2	0.481		
		0.671	C2	R3	0.573		
		0.767	C3	R4	0.591		
		0.710	C4	R5	0.627		
				R6	0.497		
				R1			-0.559

Finally, the examination of the factor covariances of both questionnaires (Tables 11 and 12) is revealing of other challenges.

Table 11

Factor Covariances for the BLTHQ

							95% Confidence Interval	
			Estimate	Std. Error	z-value	p	Lower	Upper
C	↔	O	0.287	0.100	2.880	0.004	0.092	0.482
C	↔	S	-0.560	0.076	-7.370	< .001	-0.709	-0.411
O	↔	S	-0.028	0.104	-0.271	0.786	-0.232	0.175

Table 12

Factor Covariances for the OHF-Questionnaire

							95% Confidence Interval	
			Estimate	Std. Error	z-value	p	Lower	Upper
P	↔	Q	-0.135	0.098	-1.375	0.169	-0.327	0.057
P	↔	R	-0.488	0.075	-6.486	< .001	-0.635	-0.340
Q	↔	R	0.403	0.086	4.700	< .001	0.235	0.572

In short, *p* expresses that, in both models, subjectivist items (S or Q) and objectivist items (O or P) do not correlate, while criterialist items (C or R) correlate with the two other stances in contradictory directions: criterialist items correlate positively with objectivist items in the BLTHQ we tested, but negatively in our OHF; likewise, criterialist items correlate negatively with subjectivist items in the BLTHQ, but positively in the OHF. This demonstrates that the criterialist items are problematic in both cases and could be interpreted as a strange case of "wobbling": in the BLTHQ, this "wobbling" is between criterialist and objectivist items, while in the OHF, it happens between subjectivist and criterialist items.

Discussion and conclusion

Although the results of our study seem interesting, they are also ambiguous, confirming the complexity and difficulties of assessing epistemic conceptions through a quantitative approach.

On one hand, it is interesting to have been able to extract a three-factor solution in both cases – we even tested a two-factor solution, and it did not work at all. As such, our study contradicts the results obtained by Maggioni et al. (2004) and is in agreement with the theorical tenet of a three-stance model to assess epistemic beliefs. However, the level of explanation of the total variance in both models is relatively modest, and lower than the 41% shown in the SOSCIE-study (Mierwald et al., 2017, p. 186). Moreover, although the OHF-questionnaire seems to be a slightly better fit, the BLTHQ displays a better internal consistency, in contradiction with our expectations.

On the other hand, we observed puzzling correlation patterns between the different factors in each model, that may be related to the phenomenon of "wobbling" (Maggioni et al., 2009) observed in the previous studies, but this time with a new twist. In the first questionnaire, respondents seem to "wobble" between criterialist and objectivist items, while in the second model, they "wobble" between the criterialist and the subjectivist items. In both cases, the criterialist items appear to be the stance participants 'wobbled to'. These patterns are contradictory to previous studies (Maggioni et al., 2004; Maggioni et al., 2010; Maggioni et al., 2009; Mierwald et al., 2016; Miguel-Revilla, 2022), that showed a 'wobbling' between the objectivist and subjectivist stances.

These findings could be due to different reasons to be explored in further studies.

Firstly, our results may have been affected by the differences of the sample sizes and especially the rather limited numerosity of the data related to the BLTHQ.

Secondly, another limitation may be due to the sampling of the respective groups. Our investigations were conducted among university student teachers in History, which is a sample of people with a prior interest in History. As shown in previous studies, particular interest can enhance more nuanced considerations (Stoel et al., 2022, p. 20): the fact that the participants have a particular interest in history could explain some of the wobbling towards criterialist items. Besides, the different compositions of the respective groups, regarding their age, study experience and other criteria might also have impacted our results. As shown in former studies (Miguel-Revilla et al., 2020; Nitsche, 2019; Perry, 1970), these factors tend to influence people's epistemic beliefs.

Thirdly, our results suggest that differences in the wording of the items, even small, can have a decisive impact on their interpretations by the respondents. Thus, the introduction of fewer and more nuanced statements in the OHF than in the BLTHQ may explain the contradiction between displaying a slightly better fit together with a weaker consistency. In addition, our results may also indicate hermeneutical difficulties due both to the understanding and the developmental structure/approach of the questionnaires. Being composed of complex questions that can easily be misunderstood or misinterpreted either in themselves, or due to their wording or the order in which they are presented, the developmental nature of the approach may involve that the questionnaire is primarily designed to be fully understood by respondents who are (presumably) at the third/highest stage. Participants that are (supposedly) at the first stage – because they are at the first stage and do not have the necessary prerequisites – will be more likely to

misunderstand the items belonging to stages two and three. If the same problem occurs for respondents at stage two, one may understand the lack of internal consistency and the "wobbling" observed.

These interpretations and problems suggest different avenues and solutions for future research. Further study could work towards creating a better questionnaire, by simplifying and sharpening the wording of different statements. Following on from our comparative study, one can imagine retaining the statements with the highest results from both questionnaires, and revising the statements with low results (e.g. items O9, S8, C1, P4, P5, P6, Q1 and Q6). In addition, increasing the number of items belonging to each stance could be a good idea to increase the consistency of the new tool. The next step could be to assess, qualitatively, through think-aloud individual interviews, how respondents understood the different statements of the questionnaire, to identify the abovementioned potential hermeneutical issues or issues inherent to the developmental approach of the instrument. More generally, such cognitive interviews to secure the quality of future questionnaires appear necessary, not only among experts, but also among the group of individuals researchers will be targeting for further studies. Lastly, another objective of future study could work towards creating a larger questionnaire, by integrating dimensional perspectives or prioritization scales, in order to refine the instrument and address the problem of "wobbling". Given the believed impact of epistemic beliefs on students' and teachers' practices and performance, improving the means to evaluate these beliefs remains an important goal for history education.

References

Barton, K. C., & Levstik, L. S. (2009). *Teaching history for the common good*. Routledge.

Barzilai, S., & Weinstock, M. (2015). Measuring epistemic thinking within and across topics: A scenario-based approach. *Contemporary Educational Psychology*, *42*, 141-158. https://doi.org/https://doi.org/10.1016/j.cedpsych.2015.06.006

Brownlee, J. L., Ferguson, L. E., & Ryan, M. (2017). Changing teachers' epistemic cognition: A new conceptual framework for epistemic reflexivity. *Educational Psychologist*, *52*(4), 242-252. https://doi.org/https://doi.org/10.1080/00461520.2017.1333430

Buehl, M. M., & Alexander, P. A. (2001). Beliefs about academic knowledge. *Educational Psychology Review*, *13*(4). https://doi.org/http://dx.doi.org/10.1023/A:1011917914756

Buehl, M. M., & Fives, H. (2016). The role of epistemic cognition in teacher learning and praxis. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 247-264). Routledge. https://doi.org/https://doi.org/10.4324/9781315795225

Hofer, B. K., & Bendixen, L. D. (2012). Personal epistemology: Theory, research, and future directions. In K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, & J. Sweller (Eds.), *APA Educational Psychology Handbook, Vol. 1: Theories, Constructs, and Critical Issues* (pp. 227-256). American Psychology Asssociation. https://doi.org/https://doi.org/10.1037/13273-009

Journell, W. (2017). Fake news, alternative facts, and Trump: Teaching social studies in a post-truth era. *Social Studies Journal*, *37*(1), 8-21. https://pcssonline.org/wp-content/uploads/2017/06/2017-Spring-PCSS-Journal.pdf#page=8

King, P. M., & Kitchener, K. S. (1994). *Developing Reflective Judgment: Understanding and Promoting Intellectual Growth and Critical Thinking in Adolescents and Adults*. Jossey-Bass Publishers.

King, P. M., & Kitchener, K. S. (2002). The Reflective judgment model: Twenty years of research on epistemic cognition. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 37-61). Routledge. https://doi.org/https://doi.org/10.4324/9780203424964

Kuhn, D. (1999). A developmental model of critical thinking. *Educational researcher*, *28*(2), 16-26. https://doi.org/https://doi.org/10.3102/0013189X028002016

Kuhn, D., & Weinstock, M. (2002). What is epistemological thinking and why does it matter? In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 121-144). Lawrence Erlbaum Associates. https://doi.org/https://doi.org/10.4324/9780203424964

Lee, P., & Shemilt, D. (2003). A scaffold, not a cage: Progression and progression models in history. *Teaching History*(113), 13. https://doi.org/https://www.jstor.org/stable/43259908

Lévesque, S. (2008). *Thinking historically: Educating students for the twenty-first century*. University of Toronto Press.

Maggioni, L. (2010). *Studying epistemic cognition in the history classroom: Cases of teaching and learning to think historically* [Doctoral thesis, University of Maryland]. Ann Harbor, MI. https://www.proquest.com/docview/762381299

Maggioni, L., Alexander, P., & VanSledright, B. (2004). At a crossroads? The development of epistemological beliefs and historical thinking. *European Journal of School Psychology*, *2*(1-2), 169-197.

Maggioni, L., Fox, E., & Alexander, P. A. (2010). The epistemic dimension of competence in the social sciences. *JSSE-Journal of Social Science Education*, *9*(4), 15-23. https://doi.org/https://doi.org/10.4119/jsse-538

Maggioni, L., VanSledright, B., & Alexander, P. A. (2009). Walking on the borders: A measure of epistemic cognition in history. *The Journal of Experimental Education*, *77*(3), 187-214. https://doi.org/https://doi.org/10.3200/JEXE.77.3.187-214

Mathis, C., & Parkes, R. (2020). Historical Thinking, Epistemic Cognition, and History Teacher Education. In C. W. Berg & T. M. Christou (Eds.), *The Palgrave Handbook of History and Social Studies Education* (pp. 189-212). Springer International Publishing. https://doi.org/10.1007/978-3-030-37210-1_9

Mierwald, M. (2020). *Historisches Argumentieren und epistemologische Überzeugungen: Eine Interventionsstudie zur Wirkung von Lernmaterialien im Schülerlabor*. Springer-Verlag. https://doi.org/https://link.springer.com/book/10.1007/978-3-658-29955-2

Mierwald, M., & Junius, M. (2022). Thinking aloud about epistemology in history: How do students understand the Beliefs about History Questionnaire. *Historical Encounters*, *9*(1), 35-57. https://doi.org/https://doi.org/10.52289/hej9.103

Mierwald, M., Seiffert, J., Lehmann, T., & Brauch, N. (2016). ‹ Do They Affect it All?›. Measuring Epistemological Beliefs in History Education and Their Relationship to Students' Argumentation Skills and Future History Teachers' Planning Skills. *Washington DC*.

Mierwald, M., Seiffert, J., Lehmann, T., & Brauch, N. (2017). Fragebögen auf dem Prüfstand. Ein Beitrag zur Erforschung und Weiterentwicklung eines bestehenden Instruments zur Erfassung epistemologischer Überzeugungen in der Domäne Geschichte [Questionnaires under scrutiny. A contribution to the research and further development of an existing instrument for capturing epistemological beliefs in the domain of history]. In M. Waldis & B. Ziegler (Eds.), *Geschichtsdidaktik heute: Band 8. Forschungswerkstatt Geschichtsdidaktik 15: Beiträge zur Tagung "Geschichtsdidaktik empirisch 15" [History education today: Volume 8. Research workshop on history education 15: Contributions to the conference "Empirical historical education 15"].* (pp. 177-190). hep.

Miguel-Revilla, D. (2022). What makes a testimony believable? Spanish students' conceptions about historical interpretation and the aims of history in secondary education. *Historical Encounters*, *9*(1), 101-115. https://doi.org/https://doi.org/10.52289/hej9.106

Miguel-Revilla, D., Carril-Merino, T., & Sánchez-Agustí, M. (2020). An examination of epistemic beliefs about history in initial teacher training: a comparative analysis between primary and secondary education prospective teachers. *The Journal of Experimental Education*, 1-20. https://doi.org/https://doi.org/10.1080/00220973.2020.1718059

Miguel-Revilla, D., Carril, M. T., & Sánchez-Agustí, M. (2017). Accediendo al pasado: creencias epistémicas acerca de la Historia en futuros profesores de Ciencias Sociales. *REIDICS. Revista de Investigación en Didáctica de las Ciencias Sociales*(1), 86-101. https://doi.org/https://doi.org/10.17398/2531-0968.01.86

Miguel-Revilla, D., & Fernández, J. (2017). Creencias epistémicas sobre la Geografía y la Historia en la formación inicial del profesorado de Educación Infantil y Primaria. *Didactica de La Ciencias Experimentales y Sociales*(33), 3-20. https://doi.org/http://doi.org/10.7203/dces.33.10875

Muis, K. R., Duffy, M. C., Trevors, G., Ranellucci, J., & Foy, M. (2014). What were They Thinking? Using Cognitive Interviewing to Examine the Validity of Self-Reported Epistemic Beliefs. *International Education Research*, *2*(1), 17-32.

Namamba, A., & Rao, C. (2016). Teachers' Beliefs about History and Instructional Approaches: A Survey of Secondary School Teachers in Kigoma Municipality, Tanzania. *International Journal of Education and Research*, *4*(8), 203-220. http://ijern.com/journal/2016/August-2016/15.pdf

Nitsche, M. (2019). *Beliefs von Geschichtslehrpersonen: eine Triangulationsstudie [Beliefs of history teachers: a triangulation study]*. hep verlag.

Perry, W. G. (1970). *Forms of Intellectual and Ethical Development in the College Years: A Scheme.* Holt, Rinehart, and Winston.

Royer, J. M. (2006). *The Cognitive Revolution in Educational Psychology*. IAP.

Sakki, I., & Pirttilä-Backman, A.-M. (2019). Aims in teaching history and their epistemic correlates: a study of history teachers in ten countries. *Pedagogy, Culture & Society*, *27*(1), 65-85. https://doi.org/https://doi.org/10.1080/14681366.2019.1566166

Shemilt, D. (1980). *Evaluation Study: Schools Council History 13-16 Project*. Holmes McDougall. www.schoolshistoryproject.co.uk/wp-content/uploads/2015/12/NewLookAtHistory.pdf

Stoel, G., Logtenberg, A., & Nitsche, M. (2022). Researching epistemic beliefs in history education: A review. *Historical Encounters*, *9*(1), 11-34. https://doi.org/https://doi.org/10.52289/hej9.102

Stoel, G., Logtenberg, A., Wansink, B., Huijgen, T., van Boxtel, C., & van Drie, J. (2017). Measuring epistemological beliefs in history education: An exploration of naïve and nuanced beliefs. *International Journal of Educational Research*, *83*, 120-134. https://doi.org/https://doi.org/10.1016/j.ijer.2017.03.003

Stoel, G., van Drie, J., & van Boxtel, C. (2017). The effects of explicit teaching of strategies, second-order concepts, and epistemological underpinnings on students' ability to reason causally in history. *Journal of Educational Psychology*, *109*(3), 321-337. https://doi.org/http://dx.doi.org/10.1037/edu0000143

VanSledright, B., & Limón, M. (2006). Learning and teaching social studies: A review of cognitive research in history and geography. In P. A. Alexander & P. H. Winne (Eds.), *Hanbook of Educational Psychology* (pp. 545-570). Erlbaum. https://doi.org/https://doi.org/10.4324/9780203874790

Wansink, B. G.-J., Akkerman, S., & Wubbels, T. (2016). The Certainty Paradox of student history teachers: Balancing between historical facts and interpretation. *Teaching and Teacher Education*, *56*, 94-105. https://doi.org/https://doi.org/10.1016/j.tate.2016.02.005

About the Authors

David-Alexandre Wagner is an associate professor in history and social studies at the Department of Cultural Studies and Languages, University of Stavanger, Norway, where he teaches history, historiography and history education. He is a founding member of the Future Pasts Group (FPG), that does research in History Education, Memory Studies and Cultural History. His research interests focus on combining theory and practices in history education, especially in relation to the use of visual media like films, and more recently in relation to epistemology.

Email: david.wagner@uis.no
ORCID: https://orcid.org/0000-0002-2060-3816

Ulrich Dettweiler is a professor of pedagogy at the Faculty of Arts and Education, University of Stavanger (UiS) and currently heads the faculty's doctoral programs and the university's Cognitive and Behavioral Neuroscience Lab. His research mainly focuses on cognitive and motivational aspects of Education Outside the Classroom and their biophysiological substrates. He is also interested in epistemological, methodological and ethical aspects of research.

Email: ulrich.dettweiler@uis.no
ORCID: https://orcid.org/0000-0002-8636-0298