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 and written historical reasoning: Exploring their relationship

Kristin A. Sendur[ab]
Jannet van Drie[a]
Carla van Boxtel[a]

[a]*Research Institute of Child Development and Education, University of Amsterdam, Netherlands*
[b]*Center for Individual and Academic Development, Sabanci University, Turkey*

Abstract

In this descriptive study, we investigated undergraduate students' epistemic beliefs in history and examined the relationship between students' beliefs and their performance in written historical reasoning in the context of a historical reasoning course. We measured students' expressed epistemic beliefs in history through a discipline-specific survey, which we compared with students' performance when writing a source-based historical argument. A subset of students also participated in a task-based interview to investigate more tacit epistemic beliefs related to the second-order concept, account. We found a significant correlation between students' performance in source-based argumentative writing and their epistemic beliefs regarding historical methodology. Most students' interview answers corresponded to their epistemic beliefs as indicated in the survey, but there was less correspondence between students' interviews and writing. This study demonstrates the usefulness of the epistemic beliefs survey and provides evidence that students' conceptions of the second-order concept, account, may be related to their epistemic beliefs.

Keywords

Writing, higher education, epistemic beliefs, historical reasoning

Citation

Sendur, K. A., van Drie, J., & van Boxtel, C. (2022). Epistemic beliefs and written historical reasoning: Exploring their relationship. *Historical Encounters, 9*(1), 141-158. https://doi.org/10.52289/hej9.108

Introduction

There is growing interest in epistemic beliefs and how these relate to learning in different domains. Muis, Bendixen, and Haerle (2006) have concluded that epistemic beliefs, or beliefs about knowledge and knowing, have both domain-general and domain-specific aspects. In the discipline of history, the focus of this study, this domain-specificity has been studied in different ways, including how students respond to discrepant, or conflicting, accounts of an event or topic (Barzilai & Weinstock, 2015).

One important reason to study epistemic beliefs in history is because of their potential role in how students engage in historical reasoning and source-based writing. Students' beliefs may affect how they view the construction of knowledge (Maggioni, Fox, & Alexander, 2010) and their intellectual performance (Kuhn, 2001). Without more nuanced beliefs, this could influence how students use heuristics such as sourcing when reasoning historically or how they approach a historical writing task. In a study comparing history professors and university students' approach to writing tasks in history, the professors approached the tasks from an interpretative lens that was missing from many of the students, particularly in report writing (Greene, 1994). A relationship between epistemic beliefs and written historical reasoning could help explain such findings.

Frameworks in history education have proposed that a connection between epistemic beliefs and historical reasoning may exist (Havekes, Coppen, Luttenberg, & van Boxtel, 2012; van Drie & van Boxtel, 2008; VanSledright, 2014; Wineburg, 2001). Initial findings from both qualitative and quantitative studies also provide evidence of such a relationship (Stoel, van Drie, & van Boxtel, 2017; Ioannou & Iordanou, 2020; Lee & Shemilt, 2003). Some studies have questioned the role of epistemic beliefs on historical reasoning, but have been unable to draw conclusions without a direct measure of epistemic beliefs (Monte-Sano, 2008; Reisman, 2012). These findings demonstrate that there is a need to further explore how epistemic beliefs are reflected in the way that students reason about history. Because of the difficulty of measuring epistemic beliefs, this presents a major methodological challenge. In this study, we investigate the relationship between students' epistemic beliefs in history and their written historical reasoning.

Theoretical framework

The role of epistemic beliefs in history

Epistemic beliefs have been described from multiple perspective, including developmental models (e.g. Kuhn, 2001) and independent dimensions (e.g. Hofer & Pintrich, 1997). In this study we build upon the developmental model of Kuhn and colleagues (Kuhn, 2001; Kuhn, Cheney, & Weinstock, 2000; Kuhn & Weinstock, 2002). In this model, the primary difference between the levels of beliefs is how the person reconciles objective versus subjective aspects of knowledge. At the lowest level, *realists* see claims as a copy of an objective reality, rendering critical thinking unnecessary. Those with *absolutist* beliefs hold that knowledge is objective and is generated outside of the person. Critical thinking can be utilized to determine truth. *Multiplist* believe that knowledge is an opinion constructed by the knower, and that there are not methods for determining whether one opinion is better. At the highest level are *evaluativists,* who can coordinate these objective and subjective positions. At this level, a person can use criteria to evaluate the status of knowledge.

In history, domain-specific epistemic beliefs may play a role in how students utilize second order concepts, such as evidence and historical accounts. Maggioni, VanSledright, and Alexander (2009) explored this idea when they mapped progression in the second order concept of evidence onto models of developmental domain-general epistemic beliefs. They concluded that there were similarities between Lee and Shemilt's (2003) progression model of evidence and the models of epistemic beliefs of Kuhn and Weinstock (2002) and King and Kitchener (2002). Lee and Ashby's

(2000) accounts progression model may also be linked to students' history-specific epistemic beliefs.

In the domain of history, people make judgements about claims in rival accounts. Lee and Ashby's (2000) accounts progression model describes how students move from understanding these accounts as factual representations of the past to judgements made by authors. This, in turn, affects how students discriminate between discrepant accounts. In levels one through three of their model, students see historical accounts as direct representations of the past. At level one students equate accounts with the past, similar to a *realist* in Kuhn and Weinstock's (2002) model. At level three, students allow for differences between accounts, but attribute them to "gaps in information or mistakes" (Lee & Ashby, 2000, p. 212). Level three, "*the past as determining stories*," closely resemble the *absolutist* in Kuhn and Weinstock's (2002) model since these students hold that accounts are fixed and certain. Based on Kuhn and Weinstock's (2002) model, it would be reasonable to expect students at this level in Lee and Ashby's (2000) model to use critical thinking in the discipline of history (historical methodology) as a means of verifying truth.

In levels four through six in Lee and Ashby's (2000) model, students transition to viewing the author as taking an active role in creating the account. Students at level four begin to see the authorial role in accounts, but view them as "reported in a more or less biased way" (Lee & Ashby, 2000, p. 212). This could be interpreted as conforming partly to the *multiplist* stance, as students view the author as a creator of knowledge, but flawed by authorial distortion. The accounts progression model does not go so far as to claim that students define accounts as opinions only. At level six, the highest level, students view accounts as judgements "(re-) constructed in answer to questions in accordance with criteria" (Lee & Ashby, 2000, p. 212). This closely resembles the *evaluativist* position as the student reconciles the objective and subjective dimensions of the historical account. Based on Kuhn and Weinstock's (2002) model, students may use historical methodology as a means to discriminate between rival accounts. Given these similarities, there does seem to be a link between students' domain-specific epistemic beliefs and the second order concept, account. Measuring students' epistemic beliefs and performance when using accounts may provide further evidence of the possible role of epistemic beliefs in evaluating accounts.

Measuring epistemic beliefs

Measurement is a challenge in research on epistemic beliefs. One challenge is that within a discipline, and potentially within a task, epistemic beliefs (and stances such as relativism) may be situated and contextual (Chinn, Buckland, & Samarapungavan, 2011; Chinn, Rinehart, & Buckland, 2014; Sandoval, 2014). Another challenge is that unstated, or tacit beliefs may be difficult to measure (Chinn, Buckland, & Samarapungavan, 2011; Sandoval, 2005). One method of exploring epistemic beliefs in the discipline of history may be to examine the relationship between expressed and tacit epistemic beliefs, as well as academic performance. Mixed-methods studies have been proposed as an appropriate way of examining the complex and context-specific nature of epistemic beliefs (Chinn et al., 2011; Mason, 2016).

Likert scale surveys have been used as a cost-effective method of identifying expressed epistemic beliefs since they can be administered to large groups of participants relatively easily. Such surveys have taken the form of both domain-general (Schommer, 1990) and domain-specific epistemic beliefs (Maggioni, VanSledright, & Alexander, 2009; Stoel, Logtenberg, Wansink, Huijgen, van Boxtel, & van Drie 2017). The use of Likert scale surveys has been questioned (Greene & Yu, 2014; Hofer, 2016). Hofer (2016) concludes that such scales may be appropriate for measuring very naïve beliefs, but that it may be difficult to distinguish between multiplist and evaluativist beliefs.

In this study, we administered a three-scale survey to measure expressed epistemic beliefs in history (Stoel, Logtenberg, et al., 2017). This survey was based on the model developed by Kuhn Cheney, & Weinstock (2000) and Maggion's (2010) Beliefs About History Questionnaire. First, the *nature of knowing-objective* subscale measures whether students believe that the past cannot be known because it is gone and therefor claims cannot be tested, for example, "When eyewitnesses

do not agree with each other, it is impossible to know what happened." Here, history can only be known through "true" unbiased sources. The second scale, *nature of knowledge-objective*, contains items that correspond to beliefs that history is certain and is a representation of history as it occurred (as opposed to an interpretation of the past). Students rated statements including, "When something is written in your textbook, you can be nearly certain it is true." A third scale, *historical methodology*, measures the extent to which students believe that there are criteria and procedures that can be used in history to produce knowledge. One item stated, "In history you must learn to deal with conflicting evidence." Due to issues surrounding epistemic beliefs surveys, additional measures may be useful in triangulating the data. Tasks that capture epistemic beliefs in action, such as measurements of the second order concept account, may be one option.

The evaluation of students' source-based historical writing may be an appropriate measure of academic performance to compare to their epistemic beliefs. Maggioni, Fox, & Alexander (2010) conclude that students' epistemic beliefs align with how well they are able to approach and comprehend multiple texts. Similarly, Barzilai and Eshet-Alkalai (2015) found that epistemic beliefs affected comprehension of the author's viewpoint, and in turn the use of multiple sources in a written argumentative task. Epistemic beliefs did not seem to directly affect the use of multiple sources. Therefore, it may be valuable to explore how students perform in different aspects of written historical reasoning. The heuristic sourcing, for example, may be employed for different reasons by students with a more objective versus subjective stance. Greene and Yu (2014) found that history students who prized declarative knowledge justified knowledge claims based on the veracity of that knowledge. Therefore, students with more naïve epistemic beliefs may use sourcing to determine truth, while those at higher levels may use the same heuristic to contextualize an author.

This study explores epistemic beliefs in a historical reasoning course with L2 undergraduate students. We combine a survey of students' expressed epistemic beliefs in history, a discrepant accounts task that examines tacit epistemic beliefs, and students' performance in written historical reasoning.

Research questions

In this study, we address the following research question: How do students with different epistemic beliefs, as measured by the epistemic beliefs survey and the discrepant accounts task, reason historically when writing a historical argument? Based on the theoretical framework, we would expect that students with more nuanced epistemic beliefs would display more advanced historical reasoning in their writing than those with naïve beliefs since epistemic beliefs may influence the construction of knowledge (Maggioni, Fox, & Alexander, 2010) and intellectual performance (Kuhn, 2001). From a methodological perspective, we would be interested in investigating whether there is a stronger relationship between the performance on the discrepant accounts interview and the writing than the survey and the writing because of the situated nature of the discrepant accounts and writing tasks.

Methodology

Participants

This study's data is a subset of a larger study on historical writing in English as an L2. 62 undergraduate students at a small private English-medium university in Istanbul, Turkey participated during the fall 2017 semester. See Appendix A for demographics. All students were enrolled in a Content and Language Integrated Learning (CLIL) historical reasoning course as a part of a pre-university intensive English program. Participants were non-native English speakers at the B2 level according to the Common European Framework of Reference for Languages. Students are in general 18 to 19 years old during the program.

Historical reasoning course

Students participated in a seven-week course (four hours weekly, 28 hours total) that introduced historical reasoning and argumentation using a cognitive apprenticeship model (Collins, Brown, & Holum, 1991). In the course, historical reasoning focused on argumentation (claim and evidence), historical contextualization and the heuristics source evaluation and corroboration (van Drie & van Boxtel, 2008; Wineburg, 1991). In this model, students were first introduced to a component of historical reasoning through an expert performance. Afterwards, students completed scaffolded activities before independent performance. When learning about source evaluation, for example, students first viewed an expert analysis of a primary source while taking notes using a graphic organizer. Later, students worked together to similarly analyze other primary sources with instructor support. Finally, small groups of students independently analyzed primary sources. A similar pattern was used for each aspect of historical reasoning. Students studied gladiators in the late Roman Republic and early Empire partly because it would be studied in a subsequent required history course.

During each unit students read primary and secondary sources to answer a central historical question (Monte-Sano, 2010). To facilitate reading comprehension, sources were excerpted, simplified, and presented based on recommendation by Wineburg and Martin (2009). Students used graphic organizers and guiding questions that targeted the historical reasoning concept evidence, and the heuristics, source evaluation and corroboration. Prior to reading a primary source, students identified relevant aspects of the author's background, such as the purpose in writing. Based on their notes and the central historical question, the students discussed potential ways these aspects may have influenced the author's account. While reading, students used guiding questions to address the central historical question. After reading, students considered previously read primary sources, and the extent to which the sources corroborated in answering the question.

Students wrote four *document-based question* (DBQ) essays during the course using each unit's primary and secondary sources. All questions were expositions, a one-sided argument (Coffin, 2006), and each was closely aligned with one of the central historical questions. To promote written historical reasoning in an L2, students studied aspects of language common to historical writing, such as hedging. Students used sentence stems for aspects of historical reasoning, such as source evaluation.

Throughout the course, students were challenged to integrate what they had learned about history as a discipline focused on the interpretation of the past with methodological components of historical reasoning. For example, in the first lesson, students discussed who writes about history, their varying purposes, along with the different types of available sources. By introducing students to the concept of history as interpretation and the limits of the historical record, we hoped to problematize the idea of history as truth, a commonly articulated position among our students. As a part of this lesson, we introduced methodological aspects of historical reasoning used by historians.

The course was designed by the first author and taught by twelve English language instructors using highly scripted lesson plans. Instructors had zero to six semesters experience teaching the course. Before beginning, instructors participated in a training introducing the curriculum. Extensive support material, including grading rubrics, sample essays and activity keys were provided. The course coordinator tracked lesson completion and helped ensure standardized instruction and grading. All lessons were carried out according to plan.

Data sources and analysis

There were three data sources for this study: 1) a survey of students' epistemological beliefs, 2) students' document-based question (DBQ) essays, and 3) a discrepant accounts interview.

Epistemological beliefs survey

Participants completed the epistemological beliefs survey developed by the Stoel, Logtenberg, et al. (2017). The survey had 16 questions measured on a six-point Likert scale ranging from 1 (strongly disagree) to 6 (strongly agree). The three subscales are described above.

The survey was electronically administered following the last lesson under the supervision of an instructor or research assistant. The survey was presented in simplified English and students could consult a Turkish translation. Survey administrators could clarify vocabulary, but not help students choose an answer.

The original English version of the survey was first simplified by the first author in consultation with one of the survey's authors and another postdoctoral researcher in history proficient in Dutch, English and Turkish. Two native Turkish-speaking English language instructors translated the survey into Turkish. Each translation was discussed with the same survey author to ensure fidelity to the original meaning.

Cronbach's alpha for the nature of historical knowing-objective scale was .64. The nature of historical knowledge-objective scale had an alpha of .63. Historical methodology was .92. A Kolmogorov-Smirnov test showed that the scores for the nature of knowing-objective Scale $D(62) = .14$, $p < .05$ and the historical methodology Scale $D(62) = .22$, $p < .05$, were significantly non-normal. The nature of knowledge-objective Scale $D(62) = .10$, $p = .10$, and the DBQ total score, $D(62) = .09$, $p = .20$ were normally distributed. Therefore, we conducted the non-parametric Spearman's correlation coefficient.

Document-based question

During the course, students wrote four DBQs, as described above. We analyzed the final DBQ, which was completed after instruction finished. Students answered the following question: "It is believed that many gladiators were volunteers. To what extent would it be desirable and/or undesirable for a free man to volunteer to become a gladiator?" Students had 50 minutes plus a required 10-minute planning time to write the essay using the primary and secondary sources studied in the course. The word limit was 250-300 words.

These essays were scored by the first and second author on a five-point analytical rubric (Sendur et al., 2020) and in Appendix B. This rubric was designed to assess the aspects of historical reasoning taught in the course. Cohen's Kappa from the larger dataset (This subset is 44% of the entire dataset) ranged from .66 to .82 with the claim and source evaluation categories receiving the lowest and highest scores, respectively.

Discrepant accounts interview

Following the course, students were invited to participate in interviews to further investigate their epistemic beliefs in history. Ten students (five female, five male) who volunteered and gave consent participated. See Table 1 below for participant details. Interviews were conducted by the first author and trained research assistants in English or Turkish, based on the student's preference. Interviews were audio recorded, transcribed and translated, if necessary.

In the interview, the students completed a task from Project Chata in which they read two competing accounts of the end of the Roman Empire (Lee, 2001; Lee & Ashby, 2000). All students were familiar with the topic since it is in their high school curriculum. After reading the accounts, interviewers first confirmed students' accurate comprehension of both accounts. Students then

answered the questions from the Project Chata task. See Appendix C for the questions and sample student answers. The concept of accounts was chosen because it is likely to have an epistemic aspect, as described in the theoretical framework.

Lee's (2001) description of his framework for the task contains eight subcategories subsumed into three main approaches to the task. The three major categories include: 1) students who approach the task from a *factual-based* perspective, 2) students who approach the task from the perspective that there are *multiple (factual-based) pasts*, and 3) students who used a *criterial* approach. The unit of analysis was the answer to the five questions as a whole.

The first and second author independently coded all student answers using the eight-subcategory framework. We report students' placement into one of the three main categories. Initial agreement was 60%, and all differences were resolved through discussion. As a result, each student's answer was placed into one category. The category the student was placed into was the student's predominant stance since most student displayed evidence of more than one stance within an interview (Barzilai & Eshet-Alkalai, 2015; King & Kitchener, 2002).

Results

In this section we first describe students' epistemic beliefs as measured by the survey after completing the historical reasoning course and DBQ. Next, we investigate correlations between students' beliefs and written historical reasoning. Finally, we explore the beliefs of a subset of students in the discrepant accounts interview in comparison to their survey results and written historical reasoning.

Epistemological beliefs survey

Students (N=62) scored a mean of 3.85(SD=.80) in the nature of knowing-objective scale, indicating that they partly disagreed to partly agreed that it is only possible to know about the past through unbiased sources, and a mean of 3.08(SD=.80) in the nature of knowledge-objective, indicating that they partly disagreed that knowledge in history is certain and fixed. Students scored a mean of 5.18(SD=.99) on the historical methodology scale, which indicates that most students agreed that knowledge in history is bound by disciplinary methods and criteria.

Document-based questions

Students scored a mean of 12.26(SD=3.10) on the DBQ out of 20. See Appendix A for subscores. This mean score indicates that students demonstrated an emerging level of proficiency in written historical reasoning.

Next, we investigated whether there was a correlation between students' written historical reasoning as measured by the DBQ and their epistemic beliefs as measured by the epistemic beliefs survey by calculating a Spearman's correlation coefficient. There was not a significant correlation between students' epistemic beliefs in the nature of knowing-objective scale and the DBQ score, $r_s(60) = .09$ $p > .05$ or between the nature of historical knowledge-objective scale and the DBQ score, $r_s(60) = -.04$ $p > .05$. There was a significant positive correlation between students' beliefs in the historical methodology scale and the DBQ score, $r_s(60) = .46$, $p < .01$. We also examined DBQ subscores for the historical methodology scale and found significant positive correlations for the subscores evidence, $r_s(60) = .39$, $p < .01$, source evaluation, $r_s(60) = .33$, $p < .01$, and corroboration, $r_s(60) = .29$, $p < .05$. We concluded that students' epistemic beliefs regarding historical methodology, as measured by the survey correlate with their written historical reasoning, as measured by the DBQ.

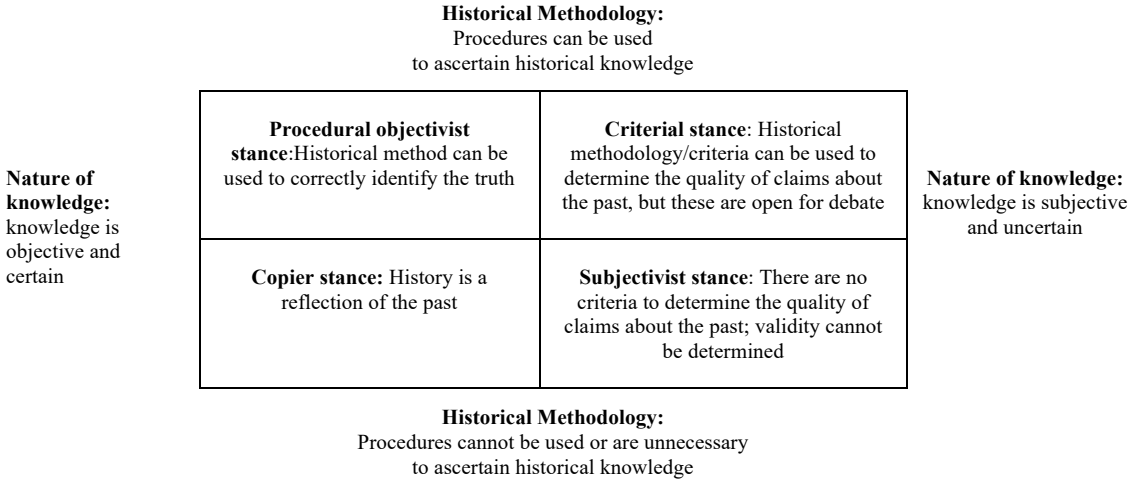
Discrepant accounts interview

An analysis of the ten students who participated in the discrepant accounts interview showed that three were classified as using a factual approach to the task, as shown in Table 1 (column Discrepant Accounts Approach), below. These students used historical methodology to assess the veracity of the accounts. Seven students used a criterial approach to determine the fall of the Roman Empire.

Triangulating data: Epistemic stance and interview

In this section we explore ten students' epistemic beliefs by triangulating their performance in the discrepant accounts interview, the epistemic beliefs survey and the DBQ. Based on the nature of knowledge-objective and historical methodology survey scales, we identified four potential epistemic stances corresponding to the level of agreement with these two scales. We did not include the nature of knowing-objective scale when identifying stances since it did not add substantially to the analysis and the stances identified in the literature correspond well to the two utilized scales (Maggioni et al., 2009; Stoel, Logtenberg, et al., 2017). We tentatively placed students into an epistemic stance corresponding to their survey responses, as shown in Figure 1. Students who completed the discrepant accounts interview fell into two stances: criterial and procedural objectivist.

FIGURE 1: Epistemic stances based on the nature of knowledge-objective and historical methodology scales from the epistemic beliefs survey



We categorized six students who had high agreement with the historical methodology scale and low agreement with the nature of knowledge-objective scale as criterial, in line with Lee's (2001) terminology. We expected these students to approach the discrepant accounts interview using historical methodology and/or criteria to determine the quality of claims about the past, but consider these claims as uncertain, similar to an evaluativist (Kuhn & Weinstock, 2002). We thus expected that they would use a criterial approach in the interview.

The second group of four students indicated strong agreement with both the historical methodology and nature of knowledge-objective scales. We classified these students as procedural objectivists. Similar to the criterial group, we expected these students to value the use of historical methods. Unlike the criterial stance, however, we expected these students to believe that historical methodology could be used to determine an objective truth, similar to an absolutist (Kuhn & Weinstock, 2002). Because of their epistemic beliefs, we expected these students to use a factual approach in the discrepant accounts interview.

TABLE 1: Scores for the epistemological beliefs survey, discrepant accounts interview and DBQ for a subset of students (n=10)

Student	Epistemological Beliefs Survey		Epistemic Stance	Discrepant Accounts Approach	DBQ Total Score
	Nature of knowledge-objective	Historical methodology			
S1	2.20	5.50	Criterial	Criterial	10
S2	3.60	5.83	Procedural objectivist	Criterial	14
S3	4.20	6.00	Procedural objectivist	Factual	17
S4	3.20	5.83	Criterial	Criterial	16
S5	4.00	5.33	Procedural objectivist	Factual	10
S6	3.20	6.00	Criterial	Criterial	14
S7	2.20	5.67	Criterial	Criterial	11
S8	3.60	5.50	Procedural objectivist	Factual	13
S9	1.80	4.50	Criterial	Criterial	15
S10	2.60	6.00	Criterial	Criterial	9

Note: DBQ scores can range from 0 to 20.

There appears to be a positive relationship between the epistemic stance as determined by the survey results and the discrepant accounts interview approach. With the exception of one student, S2, students' discrepant accounts interview approach aligned with our expectations based on the survey results. S2 was classified as a procedural objectivist based on the survey, but used a criterial approach in the interview.

Students categorized as criterial by the survey scored highly on the interview and used both interpretive criteria, such as their definition of the concept of an empire, and historical methodology, like source evaluation, to analyze the discrepant accounts. S4, for example, noted the interpretive criteria used by the authors of the accounts and the possibility of additional interpretive criteria:

> Both of them looked from different perspectives. One of them looked from more economic perspective. One of them looked from more military perspective. This question is about that every issue can be interpreted from different perspectives. (S4, interview excerpt).

These students also noted the importance of historical methodology. While some students simply noted that "historians have their own methods" (S9, interview excerpt), others explained the way they used historical methodology and interpretive decisions together:

> We can find sources that would support these two stories and we can make a comparison among those sources. We need to judge these sources in terms of their authors, contents for reliability. Then we need to research whether we can take Byzantium as the Roman Empire. That is how we can decide (S6, interview excerpt).

In contrast, students (except S2) categorized as procedural objectivists focused on historical methodology. When asked to decide how the Empire ended, these students used historical methodology as the sole criteria for determining the fall, as in this excerpt:

> We can look to writer's life story and the time period he lived. (So) we can decide it's wrong or right. (Based on) his relationships, purpose and perspective. And time period. Maybe he saw this event. And he might lived it (S3, interview excerpt).

This focus is possibly because they saw the accounts themselves as the source of knowledge and sought to evaluate the veracity of the accounts using this methodology, similar to the students in Greene and Yu's (2014) study.

Triangulating data: Epistemic stance and writing

Here we present four cases illustrating how students with differing epistemic stances perform on the DBQ. In the first two cases, the student's epistemic stance aligns with their expected performance on the DBQ, while in the second two cases it does not.

S4's criterial stance aligns with his strong DBQ performance. In his interview, he focused on how historians can approach a historical question from multiple perspectives (see above). He also connected these perspectives to historical methodology:

> Since we cannot observe the history directly, we search issues from historical sources. The sources that we use, people that we consult, our way of thinking, how we reconcile the ideas and the reliability of sources would influence the outcome. (interview excerpt)

S4 approaches his DBQ similarly. He first considers the evidence of inherent danger and social penalties for gladiators. However, in the following excerpt he ultimately reconciles this evidence with multiple sources advocating for advantageous aspects of becoming a gladiator to argue that it was a reasonable risk for those in lower (but not upper) socioeconomic classes.

> According to Kyle, gladiators' life conditions were harsh but better than poor Romans considering housing, medical attention and food opportunities. This is supported by Dunkle who agrees a career as a gladiator would attract free men because of the chance gladiators might become famous and wealthy which would buy their freedom (Dunkle 2002). Use of this evidence demonstrates that being a gladiator is a reasonable risk to get a better life for a free man. (DBQ excerpt)

S4's use of historical methodology, such as corroboration, combined with his argumentative approach result in a high score on his DBQ.

Next, we examine the case of S3, a student classified as a procedural objectivist, whose DBQ performance and epistemic stance also align. S3 received the highest DBQ score among the interviewed students. In her interview, S3 focused on the methodology of historical reasoning, as expected given her stance. When determining the date for the fall of Rome she mentioned both source evaluation and corroboration, noting "And if a two or three, four, texts are the same, dates are the same, I think it's might be true. But we can't decide, we can't decide just by looking at one." Unlike students classified as criterial, however, her use of historical methodology was to find the objective truth.

This student, in turn, wrote a DBQ demonstrating good control of the methodology of historical reasoning including well-formulated source evaluation, evidence from multiple sources, historical contextualization, and a claim. One aspect of her essay that was striking was that she included all of the evidence. This is evident in her expansive (but not entirely accurate) claim:

> Volunteering for being a gladiator is desirable because they will have citizen rights, will live better than many people and will be an object of female adoration. Also volunteering is undesirable for the reasons that gladiators will be unhealthy and people from the upper class see them in the group of shames and they separate them (DBQ excerpt).

The lack of selection in her claim and evidence may point to a lack of criteria, which is unsurprising given her epistemic stance.

The case of S1, categorized as a criterial stance, demonstrates that in some cases, the epistemic beliefs of the student do not result in a DBQ with high quality written historical reasoning. In his interview, S1 focused on his criteria for defining an empire, which he explained as "when I think

the concept of the empire, the thing that comes my mind is that people from various nations live all together." He uses this criteria as a means of evaluating his level of agreement with the two accounts. In his answer, however, he is troubled by the uncertainty of history, which he attributes partially to the reliability of his sources, concluding, "but in the end of the day the history may change in accordance with new resources. The thing that we know as true may be disproved. That's why we can't be precise."

Unlike the criterial approach in his interview, S1's DBQ reads more as a compilation of evidence than an interpretation. Notably, S1's DBQ does not indicate the use of any criteria to guide his answer. S1 does carry one feature of his interview to his DBQ: his concern about the reliability of his sources. This is visible in the extensive source evaluations he includes for each of his sources, one of which follows:

> A similar point is made by Cicero, who was a rich politician in the Roman Republic…Cicero witnessed the gladiator games and lived at the end of Republic. However, he probably sponsored the games so he might exaggerated the bravery of gladiators. Probably he was a partly reliable source (DBQ excerpt).

The end result is a DBQ that demonstrates some emerging markers of written historical reasoning, but does not match the criterial stance of his survey or discrepant accounts interview.

Finally, we profile S5, a student classified as a procedural objectivist who performed poorly on the DBQ. In her interview, S5 argued that there is a single objective history, but that it is difficult or impossible to ascertain because sources can be "made up" to fulfill a purpose. While distrustful of the sources, she still maintained that the best way to find the truth was "maybe from people who lived in those times could tell the real story like they could call eye witnesses."

Similar to S1, S5's DBQ read as a compilation of evidence and poorly formulated methodological aspects of historical reasoning that did not come together to form a coherent argument, and was in fact, at some points at odds with the available evidence. The strongest aspect of the essay was her use of source evaluation:

> Suetonius was useful as a source for gladiators' importance because he could have accessed to information that Emperor Augusts wrote and was in charge of Roman libraries which makes him a reliable source (DBQ excerpt).

Her use of sourcing, while reasonable, was offset from the evidence she used from the same source, rendering it less effective. The resulting DBQ made poor use of historical methodology, which is at odds with her epistemic stance.

Discussion and conclusions

In this study, we investigated how students with different epistemic beliefs, as measured by a domain-specific survey and a discrepant accounts interview reason historically when writing a historical argument.

In the epistemic beliefs survey, students indicated that they partly disagreed to partly agreed that it is only possible to know about the past through unbiased sources. They partly disagreed that history is fixed and certain, and agreed that knowledge in history is bound by disciplinary methods and criteria. These expressed beliefs demonstrate a somewhat nuanced view of history as a discipline. Since there is evidence of a relationship between epistemic beliefs and intellectual performance in general (Kuhn, 2001) and how students approach multiple sources in history specifically (Maggioni et al., 2010), we would expect that students' beliefs in this study would also show a similar relationship. This hypothesis was partially confirmed. Students' DBQ scores were positively correlated with the belief that history is bound by disciplinary methods and criteria, but not with the other scales. Since writing was assessed for the application of historical reasoning (i.e., disciplinary methods and criteria), this relationship is not unexpected.

A subset of students participated in a discrepant accounts interview, which explored their tacit epistemic beliefs about the second order concept account. Students' epistemic stance, as identified by the survey, largely corresponded to their approach in the discrepant accounts interview. We compared these students' epistemic stance with their DBQ performance, expecting a relationship between the two. Illustrations of students whose DBQ performance did and did not align with their epistemic stance demonstrate that this expectation is partially confirmed. Two potential reasons that students' beliefs and DBQ performance might not consistently align may be that these students view the DBQ task as non-argumentative (Greene, 1994) or that these L2 students have difficulty when writing from sources (Cumming, Lai, & Cho, 2016). Another possible reason is that the epistemic stance and the discrepant accounts interview provide information about whether students believe that there are procedures and criteria that can be used to determine the quality of claims, but do not provide information about the extent to which students are able to apply those procedures and criteria.

This study is limited by the small number of students who participated in interviews and the low reliability of two of the survey scales. Studies with a larger population and using a fully validated version of the simplified survey are needed. A more recent DBQ topic could also be used to investigate if the topic plays a role in students' written historical reasoning. While it is possible that students' knowledge of historical methodology influenced their performance, the relationship between the interview and the survey indicate that this study has likely captured aspects of students' epistemic beliefs and their influence on written historical reasoning.

This study makes a contribution towards the difficult task of measuring epistemic beliefs. Triangulating a survey of epistemic beliefs, a discrepant account interview, and written historical reasoning allowed for a richer picture of students' epistemic beliefs and supports the idea that the survey is a reasonable measure of students' epistemic beliefs in history. This is an important contribution since the use of surveys to measure epistemic beliefs is seen as problematic (Hofer, 2016). It also provides evidence that beliefs about the second order concept account described by Lee and Ashby (2000) are associated with epistemic beliefs, similar to findings related to the second order concept evidence (Maggioni et al., 2010; Maggioni et al., 2009). Future research may explore more efficient ways of using these concepts to examine epistemic beliefs to benefit from data triangulation.

Acknowledgement

The authors thank Gerhard Stoel for his support in translating the epistemic beliefs survey.

References

Barzilai, S., & Eshet-Alkalai, Y. (2015). The role of epistemic perspectives in comprehension of multiple author viewpoints. *Learning and Instruction, 36,* 86-103. https://doi.org/10.1016/j.learninstruc.2014.12.003

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/10.1016/j.cedpsych.2015.06.006

Chinn, C. A., Buckland, L. A., & Samarapungavan, A. (2011). Expanding the dimensions of epistemic cognition: Arguments from philosophy and psychology. *Educational Psychologist, 46(*3), 141-167. https://doi.org/10.1080/00461520.2011.587722

Chinn, C. A., Rinehart, R. W., & Buckland, L. A. (2014). Epistemic Cognition and Evaluating Information: Applying the AIR Model of Epistemic Cognition. In D. N. Rapp & J. L. G. Braasch (Eds.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences* (pp. 425). Cambridge: MA: MIT Press.

Coffin, C. (2006). *Historical discourse: The language of time, cause and evaluation*. London: Bloomsbury Publishing.

Collins, A., Brown, J. S., & Holum, A. (1991). Cognitive apprenticeship: Making thinking visible. *American educator, 15*(3), 6-11.

Cumming, A., Lai, C., & Cho, H. (2016). Students' writing from sources for academic purposes: A synthesis of recent research. *Journal of English for Academic Purposes, 23,* 47-58. https://doi.org/10.1016/j.jeap.2016.06.002

Greene, J. A., & Yu, S. B. (2014). Modeling and measuring epistemic cognition: A qualitative re-investigation. *Contemporary Educational Psychology, 39*(1), 12-28. https://doi.org/10.1016/j.cedpsych.2013.10.002

Greene, S. (1994). The problems of learning to think like a historian: Writing history in the culture of the classroom. *Educational Psychologist, 29*(2), 89-96. https://doi.org/10.1207/s15326985ep2902_4

Havekes, H., Coppen, P. A., Luttenberg, J., & van Boxtel, C. (2012). Knowing and doing history: A conceptual framework and pedagogy for teaching historical contextualisation. International *Journal of Historical Learning, Teaching and Research, 11*(1), 72-93.

Hofer, B. K. (2016). Epistemic cognition as a psychological construct: Advancements and challenges. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 31-50). New York: Routledge.

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research, 67*(1), 88-140. https://doi.org/10.2307/1170620

Ioannou, K. & Iordanou, K. (2020) Elementary school students' epistemic perspective and learning strategies in history. *Learning: Research and Practice, 6*(2), 150-166. https://doi.org/10.1080/23735082.2019.1591492

King, P., & Kitchener, K. (2002). The reflective judgement model: Twenty years of research on epistemic cognition. In B. K. Hofer, & P. R. Pintrich (Ed.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 37–61): Mahwah, NJ: Lawrence Erlbaum Associates.

Kuhn, D. (2001). How do People Know? *Psychological Science, 12*(1), 1-8. https://doi.org/10.1111/1467-9280.00302

Kuhn, D., Cheney, R., & Weinstock, M. (2000). The development of epistemological understanding. *Cognitive Development, 15*(3), 309-328. https://doi.org/10.1016/S0885-2014(00)00030-7

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 Publishers.

Lee, P. (2001). History in an information culture. *International Journal of Historical Learning, Teaching and Research, 1*(2), 75-98.

Lee, P., & Ashby, R. (2000). Progression in historical understanding among students ages 7-14. In P. N. Stearns, P. Seixas, & S. Wineburg (Eds.), *Knowing, teaching, and learning history* (pp. 199-222): New York University Press.

Lee, P., & Shemilt, D. (2003). A scaffold, not a cage: Progression and progression models in history. *Teaching History, 113,* 13-23. https://www.jstor.org/stable/43259908

Maggioni, L. (2010). *Studying epistemic cognition in the history classroom: Cases of teaching and learning to think historically.* [PhD Dissertation]. University of Maryland. http://hdl.handle.net/1903/10797

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/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/10.3200/JEXE.77.3.187-214

Mason, L. (2016). Psychological perspectives on measuring epistemic cognition. In N. Valanides (Ed.), *Handbook of epistemic cognition* (pp. 375-392). New York: Routledge.

Monte-Sano, C. (2008). Qualities of historical writing instruction: A comparative case study of two teachers' practices. *American Educational Research Journal, 45*(4), 1045-1079. http://www.jstor.org/stable/27667162

Monte-Sano, C. (2010). Disciplinary literacy in history: An exploration of the historical nature of adolescents' writing. *The Journal of the Learning Sciences, 19*(4), 539-568. https://doi.org/10.1080/10508406.2010.481014

Muis, K. R., Bendixen, L. D., & Haerle, F. C. (2006). Domain-generality and domain-specificity in personal epistemology research: Philosophical and empirical reflections in the development of a theoretical framework. *Educational Psychology Review, 18*(1), 3-54. https://www.jstor.org/stable/23363851

Reisman, A. (2012). Reading like a historian: A document-based history curriculum intervention in urban high schools. *Cognition and Instruction, 30*(1), 86-112. https://doi.org10.1080/07370008.2011.634081

Sandoval, W. A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. *Science Education, 89*(4), 634-656. https://doi.org/10.1002/sce.20065

Sandoval, W. A. (2014). Science education's need for a theory of epistemological development. *Science Education, 98*(3), 383-387. https://doi.org/10.1002/sce.21107

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology, 82*(3), 498-504. https://psycnet.apa.org/doi/10.1037/0022-0663.82.3.498

Sendur, K.A., van Drie, J. & van Boxtel, C. (2020). Historical reasoning in an undergraduate CLIL course: Students' progression and the role of language proficiency. *International Journal of Bilingual Education and Bilingualism,* 1-17. https://doi.org/10.1080/13670050.2020.1844136

Stoel, G., Logtenberg, A., Wansink, B., Huijgen, T., van Boxtel, C., & van Drie, J. (2017b). Measuring epistemological beliefs in history education: An exploration of naïve and nuanced beliefs. *International journal of educational research, 83,* 120-134. 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/10.1037/edu0000143

Van Drie, J., & Van Boxtel, C. (2008). Historical reasoning: Towards a framework for analyzing students' reasoning about the past. *Educational Psychology Review, 20*(2), 87-110. https://doi.org/10.1007/s10648-007-9056-1

VanSledright, B. A. (2013). *Assessing historical thinking and understanding: Innovative designs for new standards.* New York: Routledge.

Wineburg, S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology, 83*(1), 73-87. https://doi.org/10.1037/0022-0663.83.1.73

Wineburg, S. (2001). *Historical thinking and other unnatural acts: Charting the future of teaching the past* (Critical Perspectives on the Past). Philadelphia: Temple University Press.

Wineburg, S., & Martin, D. (2009). Tampering with history: Adapting primary sources for struggling readers. *Social Education, 73*(5), 212-216. Available online: https://www.socialstudies.org/social-education/73/5/tampering-history-adapting-primary-sources-struggling-readers

About the Authors

Kristin A. Sendur is a staff member at the Center for Individual and Academic Development at Sabanci University. She completed this study as a part of her doctoral research at the University of Amsterdam. Her research interests include the learning and teaching of history, with a focus on L2 students and writing.

Email: kristin.sendur@gmail.com ORCID: 0000-0003-4736-6766

Jannet van Drie is an associate professor and teacher educator at the Research Institute of Child Development and Education of the University of Amsterdam. Her main research focus is the learning and teaching of history. She is interested in historical reasoning, the role of interaction between students (groupwork), between the teacher and students (whole-class discussions), and in the role of domain-specific writing. Publications focus on themes such as effective (domain-specific) writing instruction, dialogic history teaching, and (aspects of) historical reasoning. Next to these research activities, she works as a teacher trainer in the Graduate School of Child Development and Education and she participates in the Dutch Centre for Social Studies. Together with Carla van Boxtel she developed a framework to conceptualize and analyze historical reasoning in the classroom (Van Boxtel, C. & van Drie, J. (2018). Historical Reasoning: Conceptualizations and Educational Applications. In: S. Metzger & L. Harris (Eds). International Handbook of History Teaching and Learning. pp. 149-176. Wiley-Blackwell.

Email: j.p.vandrie@uva.nl ORCID: 0000-0003-0353-8255

Carla van Boxtel is a professor of history education at the Research Institute of Child Development and Education of the University of Amsterdam. She provides leadership to the Dutch Centre for Social Studies Education. She was trained as a historian and educational scientist and explores the learning and teaching of history in and outside schools. Carla's research focuses on historical thinking and reasoning, contextualization, historical argumentation, historical empathy, historical narratives, inquiry-based learning, collaborative learning, and museum education.

Email: c.a.m.vanboxtel@uva.nl ORCID: 0000-0002-5119-121X

Appendix A: Additional Tables

Table A1
Student Gender and Intended Area of Study (N=62)

Gender		Intended Area of Study		
Male	Female	Arts and Social Sciences	Science and Engineering	Management
32	30	4	52	6

Table A2
Descriptive statistics for the DBQ subscales (N=62)

	Mean (SD)
Claim	2.23 (1.12)
Evidence	2.58 (.93)
Source Evaluation	2.32 (1.24)
Historical Contextualization	2.44 (1.24)
Corroboration	2.69 (1.30)
Total Score	12.26 (3.10)

Note: Subscores could range between 0-4, while the total score could range between 0-20.

Appendix B: Historical Reasoning Rubric

	Claim	Use of Evidence	Source Evaluation	Historical Contextualization	Corroboration
4	Presents a clear and accurate claim that adequately addresses the question.	The evidence is accurate, relevant and sufficient to support the claim & the evidence is accurately explained at least once & explicitly linked to the claim at least once.	Refers to at least 1 author by name or title & notes relevant feature(s) of the primary source (PS). Indicates potential effect of the feature on the information &/or explains the effect &/or uses the feature to further the argument. (at least 2/3)	Provides accurate and relevant historical context (HC) (temporal, spatial or social features) as support for the claim, evidence or source. The HC is elaborate and used to situate and/or further the claim or the HC is less elaborate & explicitly used to situate and/or further the claim.	Uses multiple sources to support the same point at least once & explicitly indicates an appropriate link between the sources & explains the link by noting how they are similar
3	Presents a clear and accurate claim that partially addresses the question.	The evidence is accurate, relevant and sufficient to support the claim. The evidence may be accurately explained at least once or explicitly linked to the claim at least once.	Refers to at least 1 author by name or title & notes relevant feature(s) of the PS. Indicates potential effect of the feature on the information or explains the effect or uses the feature to further the argument. (1/3)	Provides accurate and relevant historical context It may be used to implicitly situate &/or further the argument.	Uses multiple sources to support the same point at least once & explicitly indicates an appropriate link between the sources & notes that they are similar
2	Accurately restates the question or topic without directly stating a claim. May contain minor errors.	The evidence is insufficient and may contain irrelevant or inaccurate information. The evidence is explained &/or explicitly linked to the main idea at least once. The explanation or link may be inaccurate.	Refers to at least 1 author by name or title & notes relevant feature(s) of the PS. There may be an attempt to note the effect or use it to further the argument. If included, the interpretation undermines the argument or has errors.	Provides historical context that is of limited support for the argument &/or has minor inaccuracies. It is not used to situate &/or further the argument/argument &/or there are errors.	Uses multiple sources to support the same point at least once & explicitly indicates an inappropriate or unclear link between the sources
1	The main idea is difficult to discern, implied or marginally addresses the questions &/or is inconsistent with the evidence in the sources &/or the language makes the intended meaning somewhat unclear	The evidence is insufficient and may contain irrelevant or inaccurate information. The evidence is not explained & not explicitly linked to the claim &/or the evidence is primarily copy-pasted.	Refers to at least 1 author by name or title & notes irrelevant or inaccurate feature(s) of the PS. There may be an attempt to note the effect or use it to further the argument. The interpretation may have errors.	Provides historical context that is historically inaccurate &/or largely irrelevant.	Uses multiple sources to support the same point at least once & treats sources separately without explicit corroboration (can look list-like).

Appendix C

The questions from the Project Chata task (Lee, 2001; Lee & Ashby, 2000) include: 1) You just read two different explanations of why the Roman Empire fell. What might this difference mean? a) No one knows when it ended, b) It's just a matter of opinion when it ended, c) There was no one single time when it ended, and d) One of the stories must be wrong about when it ended. They were also asked: 2) Is your choice what you really thought? 3) How could we decide when the Empire ended? 4) Are these two dates the only possible times for the end of the Empire? And 5) Do the differences between the stories matter?

TABLE C1. Main categories, subcategories and sample student answers for the Fall of Rome task

Main Category	**Subcategory**	**Student Answers (excerpt)**
Factual	1. Semantics only	No participants in this study took this approach.
	2. One story is wrong, the end is known	No participants in this study took this approach.
	3. The end is unknowable	No participants in this study took this approach.
	4. The end is knowable, but contingently unknown	I am not sure about how can we decide (which account is right) because maybe two of them are wrong. Maybe one is true but maybe from people who lived in those times could tell the real story like they could call eye witnesses maybe but I don't think we can be 100% sure about the story (S5).
Multiple Past	5. The end is multiple	No participants in this study took this approach.
Criterial	6. The end is criterial (implicit criterion)	I don't think there is another option (for an ending date). We should examine it actually…(We could find out because) historians have their own methods. It can be looked whether different historians say different things. It was exist in the establishment of Ottoman Empire. Some was saying it is 1299. One historians I don't remember the name was saying it is 1302. But, only he says that. According to what he says it I don't know. I don't know the topic that much, so I don't know (S9).
	7. The end is criterial (one explicit criterion)	When I think the concept of the empire, the thing that comes my mind is that people from various nations live all together. In the first story the West ended. The Roman Empire was shown as it ended but many people from different nations and different cultures remain living together in Eastern Empire. That's why I thought the idea of living together with different nations should be disappeared to talk about the collapse of an empire (S1).
	8. The end is criterial (alternative criteria)	So, in my opinion, I think when the capital was captured with when the people of the Empire...Not the Emperor, because the people of the Empire… they… when you capture the city, the nation's when it ended… not the emperor because the nation can improve the culture, and continue, and can teach the subsequent generations to live on. Because you can take the emperor, but the nation will keep going, nation will choose another emperor, it can change… but if you change the nation, it's another empire...(S7).

Note: Main category and subcategory titles are from Lee (2001) and Lee and Ashby (2000). Student answers are from this study's participants.