



UvA-DARE (Digital Academic Repository)

Connecting past, present and future

The enhancement of the relevance of history for students

van Straaten, D.

Publication date

2018

Document Version

Other version

License

Other

[Link to publication](#)

Citation for published version (APA):

van Straaten, D. (2018). *Connecting past, present and future: The enhancement of the relevance of history for students*.

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Chapter 5

FOSTERING STUDENTS' APPRAISALS OF THE RELEVANCE OF HISTORY BY COMPARING ANALOGOUS CASES OF AN ENDURING HUMAN ISSUE: A QUASI-EXPERIMENTAL STUDY⁴

Although history standards generally aim at developing historical consciousness among secondary school students, there is not much research-based knowledge to support making connections between the past, the present and the future in history teaching. This study examines the effects of teaching analogous cases of an enduring human issue in two experimental conditions: one in which grade 10-12 students ($n = 460$) were actively encouraged to compare cases and to draw analogies with the present and one in which students studied cases without making comparisons or drawing analogies with the present ($n = 273$). Set against the results of a group of students who followed the usual history curriculum ($n = 289$), multilevel regression analyses on the collected data revealed that both experimental conditions positively affected students' appraisals of the relevance of history, more so in the 'case-comparison' condition than in the 'separate-case' condition. Students in the case-comparison condition also deemed the lesson course more valuable and experienced less difficulty with the applied pedagogical approach than students in the separate-case condition. Case comparison did not negatively affect the acquisition of historical factual knowledge. Implications for further research are discussed.

5.1 Introduction

Developing historical consciousness is an important rationale for history as a school subject in many Western countries (e.g. DFE, 2013; NHCS, 1996; SLO, 2016; VGD, 2006). By connecting the history of mankind to their own personal lives, students should obtain a deeper understanding of today's and tomorrow's world, get a sense of their own historicity and expand their 'space of experience' thus influencing their 'horizon of expectation' (Koselleck, 2004, p. 255). As a rule, history standards cover the history of mankind from prehistoric to modern times, usually outlined in chronologically ordered

⁴ Van Straaten, D., Wilschut, A., Oostdam, R., & Fukkink, R. Fostering students' appraisals of the relevance of history by comparing analogous cases of an enduring human issue: a quasi-experimental study. Submitted to *Cognition and Instruction*, 16 January 2018 (first review received, revision in progress).

topics. However, when it comes to connecting the past to the present and the future, standards offer little guidance.

There are reasons to assume that students do not use the past as a tool for orientation on the present and future of their own accord. In many Western countries they have unarticulated views on the purposes and benefits of studying the past (e.g. Angvik & Von Borries, 1997; Biddulph & Adey, 2003; Harris & Reynolds, 2014; Haydn & Harris, 2010; Van Sledright, 1997; Zhao & Hoge, 2005). Students tend to see the past as ‘fixed’, as a closed entity of given dates and facts about a world ‘out there’ that bears little relation with the ‘real’ world; they have difficulty understanding that history is about constructing narratives about the past that serve contemporary needs and interests (e.g. Barton, 2008; Lee, 2005; Maggioni, Alexander, & VanSledright, 2004; Shemilt, 2009; Stoel, Logtenberg, Wansink, Huijgen, Van Boxtel, & Van Drie 2017). History curricula usually focus on historical topics as aims in themselves without drawing analogies with the present or referring to ‘big pictures’, thus possibly thwarting students’ ability to discern patterns of change and continuity between past and present times (Blow, 2009; Carroll, 2016; Foster, Ashby, & Lee, 2008).

Given the fact that students are not inclined to make connections between the past, present and future spontaneously, they need to be supported. In earlier work, we distinguished three pedagogical approaches for this to be achieved: (1) teaching with longitudinal lines describing long-term political, socio-economical or cultural developments, for example, the emergence of national states; (2) teaching with analogies between the past and the present, for example, an analogy between the Roman Empire and the European Union; and (3) teaching with enduring human issues (i.e. issues shared by humans of all times because they are essential to human existence, such as religious beliefs, government, trade, food and sickness) (Van Straaten, Wilschut, & Oostdam, 2016).

The present quasi-experimental study investigates the merits of history teaching that combines the second and third pedagogical approach by means of a lesson intervention conducted in Dutch senior secondary education. Our main assumption is that these approaches foster students’ abilities to use knowledge of the past in reflections about present-day affairs, thus positively affecting their appraisals of the relevance of history.

In addition, as this type of teaching is an innovative practice in Dutch history education, we investigated students' learning experiences and situational interest.

5.2 Learning through analogical reasoning

The design of this study relies on cognitive psychological research on analogical reasoning and case comparison learning. In general, case comparison activities lead to better learning outcomes than more traditional forms of instruction, such as lecturing and reading (Alfieri, Nokes-Malch, & Schunn, 2013). Comparing cases simultaneously appears to be more effective than studying cases sequentially (one after the other) without making comparisons (Alfieri et al., 2013; Gentner, Loewenstein & Thompson, 2003). In a 'sequential condition', students are not inclined to make comparisons (Gentner, Loewenstein, & Thompson, 2003; Rittle-Johnson & Star, 2007), and if they do, they focus on surface feature similarity whereas case comparison learners give much more weight to structural features, resulting in deeper conceptual understanding (Alfieri et al., 2013; Cummins, 1992).

Case comparison activities need to be accompanied by supportive cues as students, especially novices, often fail to detect structural features underlying similar cases of their own accord. Highlighting analogous features in direct instruction is helpful to students and makes case comparison activities more effective (Gentner, Loewenstein, & Thompson, 2003; Holyoak & Koh, 1987; Richland, Zur, & Holyoak, 2007). Other effective cues are: (1) searching for similarities instead of searching for both similarities and differences (Alfieri et al., 2013); (2) testing learners immediately after comparison activities (Alfieri et al., 2013); (3) using visual or schematic representations of key features (Bulgren, Deshler, Schumaker, & Lenz, 2000; Richland, & McDonough, 2010); and (4) modeling or scaffolding case comparison tasks into step-by-step learning activities (Richland, Zur, & Holyoak, 2007). Scaffolding case comparison activities may be accomplished in several diverse modes. For this study, we used the first three stages of the 'guided analogy training' model developed by Gentner, Loewenstein, & Thompson (2003): (1) introducing key principles and key concepts; (2) using a case to demonstrate the principles; (3) applying the principles on unknown cases in a comparison with the first case.

5.3 The use of analogies and the discipline of history

Thus far, cognitive psychological studies on analogical, case-based reasoning in educational contexts have mainly been conducted in the fields of mathematics and natural sciences. In these subjects, the foundation of analogical reasoning is that similar actions in similar cases will have similar effects. This may well explain why there is a lack of research on the use of analogies in history teaching (Myson, 2006). History teachers do not eschew drawing analogies, on the contrary, they use analogies quite often and there are some inspiring classroom examples of this type of teaching (e.g. Boix-Mansilla, 2000; Laffin & Wilson, 2005; Myson, 2006; Rollett, 2010). In many cases, something mundane from the present is being used to explain something similar from the past or historical events that bear strong similarities are being compared (Ata, 2009; McCarthy Young & Leinhardt, 2000). There are no data on the frequency of historical analogies being used to reflect on present-day affairs, but it seems plausible to assume this is not a regular practice, given the focus on memorizing facts in many history lessons. Moreover, this approach may seem at odds with the nature of the discipline, because history is essentially concerned with the ‘otherness’ of the past, emphasizing differences and not similarities with the present. Contingency plays an important role in historical explanation and the application of a ‘covering law model’ in this context is arguably problematic because there are no general laws which may predict human behavior (Jonker, 2001; Lorenz, 1998; Munslow, 2006).

From an academic point of view, therefore, it may seem dubious to use historical analogies for orientation on the present and future. Historical analogies may turn out to be simplistic, politically biased, false or anachronistic (Leira, 2017; Miller, 2016; Mumford, 2015). All of this is probably true for academic history, which aims at establishing historical truths as detailed as possible, although historical scholarship itself is part of a historical culture bridging past and present realities as a matter of course (Rüsen, 2017). Teaching history, however, has a more pragmatic aim, for which analogies may be quite fruitful, because they may stimulate discussion, exchange of arguments and assent (Kornprobst, 2007).

5.4 Using analogies as a relevance tool

The essence of education is preparing students for societal participation and developing their personal identity (Biesta, 2010). Studying the past should, therefore, be explicitly linked to students' lives and the society of which they are part. In earlier work, we used the concept of 'relevance of history' in this respect, which we defined as 'allowing students to recognize and experience what history has to do with themselves, with today's society and their general understanding of human existence' (Van Straaten, Wilschut, & Oostdam, 2016). We distinguished three objectives for relevant history teaching: (1) building a personal identity: seeing oneself as an individual with a personal past and developing one's own values, opinions and ideals vis-à-vis those of the historically shaped communities to which one belongs, (2) becoming a citizen: understanding the origins of contemporary affairs and developing well-substantiated views in order to function as a citizen in society, and (3) understanding the human condition: becoming aware of the temporal dimension of the human existence and supplementing one's experiences with past approaches to human issues (Van Straaten, Wilschut, & Oostdam, 2016).

Drawing analogies in the context of an enduring human issue can be a useful tool for pursuing these relevance objectives. It can help students acquire a more abstract understanding of lesson content as it involves higher order thinking skills, such as generalizing, categorizing and inferring (Richland & Simms, 2015). Through the drawing of analogies, students 'decontextualize' specific information into generally applicable principles and concepts, facilitating knowledge transfer to new examples that share underlying structural characteristics but differ in specific characteristics (Alfieri et al., 2013; Gentner, Loewenstein, & Thompson, 2003; Salomon & Perkins, 1989). All of this corresponds to research literature showing the benefits of conceptual frameworks and concept-based instruction for the teaching and learning of history (e.g. Stern, 2010; Lee, 2005; Thornton & Barton, 2010; Twyman, McCleery, & Tindal, 2006).

Using historical phenomena to reflect on analogous contemporary phenomena puts students into a position to generate insights which may have value beyond school. Recognising the utility of classroom tasks in terms of applicability in 'real life' is what encourages students to learn and what they find important in valuing the relevance of school subjects (Brophy, 1999; Eccles, 2004; Frymier & Shulman, 1995; Martin, 2003;

Muddiman & Frymier, 2009; Pintrich, 2003). Meaning making and content relevance are also important stimuli for enhancing ‘situational interest’ (i.e. the interest experienced in a particular moment emanating from environmental factors such as the clarity of tasks, the perceived value of information or the coherence and vividness of texts) (Harackiewicz, Smith, & Priniski, 2016; Schraw, Flowerday, & Leman, 2001).

5.5 Enduring human issues and the history curriculum

There have been a number of proposals for designing history curricula organized around enduring human issues exemplified by analogous cases from different periods (e.g. Barton & Levstik, 2011; Brush & Saye, 2014; Carroll, 2016; Grant & Gradwell, 2010; Hunt, 2000; McTighe & Wiggins, 2013). In many designs, enduring issues are embedded in existing curricula by means of selecting topics that incidentally seem suitable to move beyond factual historical content. Obenchain, Orr and Davis (2011), for example, apply the question ‘should liberty be limited’ to standard units such as the American Revolution (‘Was it appropriate for the Sons of Liberty to use their power to destroy property in the Boston Tea Party?’), the American Civil War (‘Was the North justified in limiting the liberties of Southern property owners?’) and the Vietnam War (‘During the Vietnam War, should the liberties of press, speech, and protest have been limited?’) (p. 193). Because essential questions are assigned to topics appearing in the curriculum for their own sake, application may often take place in very specific and very diverse contexts which must be well understood for students to be able to grasp and elaborate on the essence of the issues at stake. Moreover, as topics in existing curricula are being taught in different grades over a relatively long period of time, using analogies and abstracting generally applicable knowledge are less obvious.

For this study, therefore, we selected topics that were specifically suitable for addressing essential questions about an enduring human issue instead of embedding questions in an extant curriculum. These topics were taught sequentially in a short time span to facilitate comparison activities and the drawing of analogies between past and present. The assumption here is that comparison activities allow students to study the past in meaningful ways and consequently have a positive effect on their appraisals of the relevance of history.

5.6 The present study

A lesson unit was designed for two experimental conditions: the *case-comparison condition*, in which students discussed essential questions concerning an enduring human issue by means of comparing cases from different periods and drawing analogies with the present; and the *separate-case condition*, in which the same historical cases were taught sequentially (one at the time) without discussing essential questions, making mutual comparisons and drawing analogies with the present. We tested the extent to which case-based history teaching about an enduring human issue in both experimental conditions affected students' (1) appraisals of the relevance of history, (2) situational interest, (3) opinions about the complexity of this type of history teaching and (4) acquisition of subject matter knowledge. 'Relevance of history' corresponds to the objectives of relevant history teaching as described above (i.e. building a personal identity, becoming a citizen and understanding the human condition). 'Situational interest' refers to the way students experienced the lesson unit in terms of attention, engagement, enjoyment and value (Linnenbrink-Garcia et al., 2010).

5.7 Hypotheses

We expected that:

- (1) students' appraisals of the relevance of history are positively affected to a greater extent in the case-comparison condition than in the separate-case condition.
- (2) students' situational interest is positively affected to a greater extent in the case-comparison condition than in the separate-case condition.
- (3) students in the case-comparison condition considered the applied approach (i.e. studying cases from different periods in the context of an enduring human issue) less problematic than students in the separate-case condition.
- (4) there are no differences between the case-comparison and the separate-case condition in terms of acquisition of subject matter knowledge.

In accordance with our theoretical framework, we assumed that students in the case-comparison condition would generate generic knowledge applicable in 'real life', allowing them to experience the value of history (hypothesis 1). The case-comparison condition included stimuli for situational interest that were lacking in the separate-case

condition, such as pursuing content relevance by seeking connections with the present (hypothesis 2). Making comparisons using essential questions would enable students to envisage the historical cases in a comprehensive framework instead of seeing them as isolated events in particular historical contexts (hypothesis 3). The learning effectiveness of comparison activities gave reason to assume that students in the case-comparison condition would not underperform in terms of factual knowledge acquisition, even though a considerable part of their study time was spent on past-present analogies whereas students in the separate-case condition focused exclusively on learning historical facts and practicing historical skills (hypothesis 4).

5.8 Method

5.8.1 Study design

A pre-/posttest design with a comparison group was used to evaluate effects of the case-comparison and separate-case conditions as predicted by hypothesis 1 (see Table 5.1). To avoid confusion with the case-comparison group, we have called our comparison group ‘non-treatment group’. This group followed the regular history curriculum and completed the questionnaire concerned with a time interval between pre- and posttests as long as the average duration between pre- and posttest in the experimental conditions (i.e. five weeks). Measures in the non-treatment group were only carried out for hypothesis 1, as the other hypotheses relate to the lesson intervention in which this group had not participated. The non-treatment group also took a historical knowledge test prior to the intervention in order to examine equivalence with the treatment groups.

The outcomes of the experimental groups were not mutually compared, but independently with the outcomes of the non-treatment group. Because the tested pedagogical approaches were innovative in Dutch history education, it seemed to make sense to examine their effects on students’ history relevance perceptions compared to effects of ‘usual’ history teaching.

Table 5.1 Design and measures of the study.

	<i>Hypo-thesis</i>	<i>Pre/post</i>	<i>Case-comparis on group</i>	<i>Separate-case group</i>	<i>Non-treatment group</i>
Intervention (6 lessons)			X	X	
Measures:					
Historical Knowledge		Pre	X	X	X
History Relevance	1	Pre/post	X	X	X
Situational Interest	2	Post	X	X	
Pedagogical Approach	3	Post	X	X	
Lesson Content Knowledge	4	Post	X	X	

5.8.2 Educational context

The study was conducted in the two highest tracks of Dutch senior secondary education: middle level general secondary education (HAVO) and pre-university secondary education (VWO). The history curriculum in these tracks is based on frame of reference knowledge organized around ten eras and their characteristic features, starting with the ‘era of hunters and farmers’ and ending with the ‘era of television and computer’ (Wilschut, 2009; 2015). The ten eras are first introduced in junior secondary education and subsequently studied on a more profound level in senior secondary education. History is not a compulsory subject in senior secondary education, which means that this study’s participants had opted for a curriculum of which history was a part.

It is important to note that (1) the ten eras and their features are taught as independent topics in chronological order without comparing them in the context of enduring issues, and (2) ‘the use of history’ – a component of the Norwegian and Swedish curriculum (Nordgren, 2016) – does not appear in the Dutch curriculum. This means that, previously to our intervention, participants had not been taught any of the objectives of relevance of history underlying the constructs of the questionnaire we used to gauge students’ perceptions on this matter.

5.8.3 Participants

In the original sample, participants were 1236 grade 10 to 12 senior secondary students from 24 secondary schools located in nine out of the twelve Dutch provinces. Students who did not complete both questionnaires of the pre- and posttest measurements were excluded ($n = 214$). This resulted in a final sample of 1022 students, of whom 460 participated in the case-comparison group, 273 in the separate-case group and 289 in the non-treatment group. Table 5.2 shows the sample specifics.

Table 5.2 Participating students for each research condition and educational track.

	<i>Case-comparison condition (n = 460)</i>	<i>Separate-case condition (n = 273)</i>	<i>Non-treatment group (n = 289)</i>
Middle level general education (HAVO)	266 (57.9%)	170 (62.3%)	184 (63.7%)
Pre-university education (VWO)	194 (42.1%)	103 (37.7%)	105 (36.3%)
Mean age	16.65 ($SD = 1.01$)	16.50 ($SD = 0.95$)	16.07 ($SD = 1.03$)
Gender: female	51.1%	57.9%	51.2%

Thirty teachers participated: 22 in the experimental conditions and 8 in the non-treatment condition. As the non-treatment group was not involved in the intervention, no specifics are reported about the teachers, whose only job was to teach as usual and administer questionnaires. Teachers participating in the two experimental conditions were recruited through an announcement in a Dutch history teaching journal and by soliciting MA-level teachers graduated from our own university. They were requested to participate in both conditions in order to ensure group similarity, which was not possible in all cases, resulting in 11 taking part in the case-comparison condition, 4 in the separate-case condition and 7 in both conditions. Teaching experience of the teachers in the case-comparison condition varied from 7 to 35 years ($M = 16.50$; $SD = 8.63$); 44% were male, 56% female. Teaching experience in the separate-case condition ranged from 5 to 30 ($M = 16.27$; $SD = 8.71$); 64% were male, 36% female. Hence, all teachers in both conditions had at least five years of teaching experience.

5.8.4 Lesson unit design

Subject of the lesson unit was the enduring human issue of migration and integration. Five refugee groups from early modern times onwards were used as exemplary cases: (1) Protestants leaving the Catholic southern Netherlands for the Protestant north during the Dutch Revolt (16th century); (2) Persecuted Jews seeking refuge in the Dutch Republic (17th century); (3) Belgians fleeing from First World War violence towards the neutral Netherlands; (4) German ‘Heimatvertriebene’ being expelled from former German territory in Poland, Russia and Czechoslovakia after the Second World War; (5) Cubans leaving for the US after the communist takeover by Fidel Castro in 1959. The unit contained six lessons: one introduction lesson and one lesson for each refugee case. Pre- and post-measurements required another two lessons. To ensure that all students were equally informed, two authors – both historians, history teacher educators and history textbook writers – documented the refugee cases using sources and writing texts themselves. Each refugee lesson referred to exactly the same content subject matter in both conditions. Historical sources, texts and illustrations were identical, except for the assignments which differed in accordance with the divergent lesson goals of the two conditions (see Appendices D and E). Lesson materials were bundled into a booklet in which students noted their answers.

Table 5.3 clarifies the differences between the courses in both experimental conditions. Learning activities in the case-comparison condition were supported by modelling according to the guided analogy training of Gentner, Loewenstein and Thompson (2003) and by using a framework of key questions and concepts drawn from academic literature on migration and integration (e.g. Mavroudi & Nagel, 2016; Obdeijn & Stover, 2008) (see Appendix F). In the introduction lesson, the framework was explained by the teachers and used by students for analysing current refugee issues in the Netherlands. Students subsequently employed the framework for comparing the refugee cases and drawing analogies between these cases and present-day refugee issues. Each lesson ended with students discussing in pairs differences and similarities between the past and the present.

Students in the separate-case condition studied the historical refugee cases to deepen their knowledge about the Dutch Revolt, the Dutch Republic in the 17th century, the World Wars and the Cold War, all of which belong to the prescribed characteristic

features of the framework of ten eras. In the introduction lesson, students contextualized sources related to these topics and practised historical skills. These two assignment types – common in the history teaching students were familiar with – were consequently applied in the next five lessons about refugee groups in the past.

Table 5.3 Design of the lesson units in the experimental conditions.

	<i>Case-comparison condition</i>	<i>Separate-case condition</i>
Aim	Comparing refugee cases from the past and using them to reflect on present-day refugee issues	Learning about refugee cases from the past to deepen knowledge of general topics related to these cases
Lesson 1	Demonstrating a key questions framework for case comparison and drawing analogies with the present	Contextualizing primary sources using knowledge of general topics related to the refugee cases
	Applying the framework to current refugee issues	Practicing historical skills using the primary sources
Lesson 2-6	Studying five historical refugee cases using the same documents and images as in the separate-case condition	Studying five historical refugee cases using the same documents and images as in the case-comparison condition
Assignment types	<p>1 Analyzing and comparing refugee cases using the key questions framework.</p> <p>2 Considering present-day refugee issues by drawing analogies with the past cases.</p> <p>For example:</p> <p>[Source: Authorities in Miami issued an informative movie to reduce growing concerns among the population about the arrival of large quantities of Cuban refugees] Assignments:</p> <p>1 Which questions and concepts from the key questions framework can you relate to the concerns of the people of Miami? Choose two.</p> <p>2 The government appealed to American history to reassure the population. Nowadays, could the Dutch government refer to immigration in Dutch history for this purpose? Explain your opinion.</p>	<p>1 Contextualize primary sources using general historical knowledge.</p> <p>2 Practicing historical skills using these primary sources.</p> <p>For example:</p> <p>[Source: Authorities in Miami issued an informative movie to reduce growing concerns among the population about the arrival of large quantities of Cuban refugees] Assignments:</p> <p>1 A quote from the movie: "<i>The whole world sees the fate of the Cuban refugees.</i>" Why was it important for the US to treat the Cubans properly in the light of the Cold War?</p> <p>2 Did the movie emphasize change or continuity in American history? Explain why the authorities did this. What interest did they have?</p>

Each refugee case lesson lasted 50 minutes and had the same build-up in both conditions. The teacher introduced the topic using a digital presentation – identical for both conditions – displaying elementary facts and dates (10 minutes). Students then studied the documents and made assignments (25 minutes). The lessons ended with a plenary exchange of students’ written answers (15 minutes).

5.8.5 Teacher preparation and treatment integrity

The teachers were informed about the aims and methods of the unit in a three-hour meeting. Three absentees were personally instructed. All teachers received a guide describing goals and procedures, providing model answers and historical background information. Teachers in the separate-case condition were instructed to avoid comparing cases in order to keep a clear distinction with the case-comparison condition. During the meeting, the teachers indicated that goals and methods of the lesson intervention were clear to them.

The first author observed six lessons (three in each experimental condition) and interviewed 6 teachers and 32 students (in dyads) to collect experiences and opinions. The teachers evaluated the lesson intervention by filling out an online questionnaire. They reported their satisfaction with the quality of course materials. Students had completed the questionnaires without irregularities, the lessons had proceeded according to plan and students’ work ethic had been as usual. All teachers had completed the lesson unit, with a number of them reporting tight time schedules. Data from the interviews and the lesson observations did not point at serious deviations from lesson protocols.

5.8.6 Measures

Historical Knowledge. All three groups took a historical knowledge test at the start of the intervention in order to examine group equivalence and to control for effects of differences in knowledge levels on outcome measures (see Appendix G). We designed a pencil-and-paper test with 40 true/false statements about general historical topics related to the five refugee examples (e.g. Dutch revolt, World Wars). Item examples were: ‘Luther was pardoned by the Holy Roman Emperor at the Diet of Worms’, ‘Characteristic of 17th century Dutch painting were scenes of military battles and royal life’. ‘Russia pulled out of the First World War after the Bolshevik revolution led by

Lenin', 'South and North Vietnam were reunited under a communist government in the 1970s'. Cronbach's α was .66, indicating adequate internal consistency for this broad knowledge domain.

History Relevance. The Relevance of History Measurement Scale (RHMS) was used to gauge students' appraisals of the relevance of history (Appendix A). The RHMS is a validated closed format questionnaire designed to measure student beliefs about the value of history in view of building a personal identity, becoming a citizen and understanding the human condition (Van Straaten, Wilschut, & Oostdam, 2018). Item examples in the order of these relevance strands were: 'History helps me to get to know myself better'; 'History is of little use if you want to understand the news'; 'History enables you to imagine what will happen in the future'. The RHMS comprises 24 items each with a 6-point Likert scale (1: totally disagree; 2: disagree; 3: disagree a little; 4: agree a little; 5: agree; 6: totally agree). The reliability of the RHMS subscales was good with α -values ranging from .80 to .90. For the full scale, α was .92 for the pretest and .94 for the posttest.

Situational Interest. Students' situational interest (SI) was measured by means of a 12-item questionnaire using a 5-point Likert scale (1: totally disagree; 2: disagree; 3: neutral; 4: agree; 5: totally agree) based on an instrument designed by Linnenbrink-Garcia et al. (2010) (see Appendix H). This questionnaire measures three aspects of situational interest in classroom settings: (1) the extent to which a course grabs students' attention (triggered-SI); (2) the extent to which a course itself is pleasurable and engaging (maintained-SI-feeling); and (3) the extent to which a course is deemed important and valuable (maintained-SI-value). The original items were translated and 'in the lessons about refugees' was added to each of them to avoid students having other history classes in mind. Sample items for the three SI-aspects were respectively: 'In the lessons about refugees we did things that grabbed my attention', 'I'm excited about what we learned about refugees in history class', 'What we have learned about refugees in history class can be applied to real life'. Reliability of the three SI-subcales was good (respectively $\alpha = .82, .89$ and $.80$) Whole scale α was $.93$.

Pedagogical Approach. Students' opinions about the complexity of the applied case-comparison approach were measured by means of a self-designed 3-item questionnaire with a 5-point Likert scale (1: totally disagree; 2: disagree; 3: neutral; 4: agree; 5: totally

agree) (see Appendix I). Items were: ‘Lessons about topics from different times are confusing’, ‘Teaching a theme with topics from different times (like in the refugee lessons) is more difficult than the history teaching we are used to’ and ‘In the refugee lessons, there were so many different topics that it was difficult to understand them.’ The Pedagogical Approach scale had sufficient reliability ($\alpha = .71$).

Lesson Content Knowledge. Students’ knowledge of the five refugee cases was measured with a self-designed test comprising 30 true/false items (6 per case) (see Appendix J). Item examples are: ‘Cities in the Dutch Republic offered refugees from the Southern provinces favorable settlement conditions’, ‘About 1 million Belgian refugees arrived in the Netherlands in 1914’ and ‘Among the people of Florida there were hardly any concerns about the influx of Cuban refugees’. Reliability of the Lesson Content Knowledge scale was sufficient ($\alpha = .71$).

5.8.7 Data analysis

Preliminary analyses revealed no serious violation of the assumption of normality. Levene’s tests showed equal variance across groups for all outcome measures except for SI-total, triggered-SI and Lesson Content Knowledge. Three univariate outliers were detected. Removing them did not result in significant outcome differences, so all cases were retained; no transformations were applied.

The equivalence of the three groups at the start of the experiment was demonstrated through multilevel regression analyses on pretest outcomes for Historical Knowledge and History Relevance. Regarding their general knowledge of history, no differences were detected between the case-comparison group and the non-treatment group, $z = 0,190, p$ value (two-sided) = .849. No differences were found either between the separate-case group and the non-treatment group, $z = 0,195, p$ value (two-sided) = .849. With regard to students’ appraisals of the relevance of history, there were no differences neither between the case-comparison and the non-treatment group, $z = -0,386, p$ value (two-sided) = .700, nor between the separate-case group and the non-treatment group, $z = -0,022, p$ value (two-sided) = .982. These results automatically rule out the possibility of ‘historical knowledge’ and ‘history relevance’ differences between the case-comparison and separate-case conditions.

To check for possible effects of (differential) attrition between conditions, a MANOVA was conducted with History Relevance pretest scores as dependent variables and attrition as independent variable. No significant differences were found between students who discontinued between the pretest and posttest (these cases were removed from the sample) and students who completed all questionnaires, Wilks' $\lambda = .981$, $F(20, 3862) = 1.14$, $p = .308$. No significant interaction effect between attrition and condition was found either, Wilks' $\lambda = .997$, $F(4, 1168) = .896$, $p = .469$.

Taking into account the hierarchical data structure, with students (level 1) nested in classes (level 2), multilevel regression analysis was applied using MLwiN 2.20 (Rasbash, Steele, Browne, & Goldstein, 2009). Intra-class correlations at class level for scale measures ranged from .07 to .20, calling for multilevel analysis (Snijders & Bosker, 2012). School was not added as level because preliminary analyses rendered zero to very small intra-class correlations at school level ($\rho \leq .05$) and the number of schools was too small (Maas & Hox, 2005).

Multilevel analyses were conducted for each outcome measure: Relevance (total), Relevance-identity, Relevance-citizenship, Relevance-human condition; Situational Interest (total), Maintained-SI-Value, Maintained-SI-Feeling, Triggered-SI; Pedagogical Approach; Lesson Content Knowledge. The two experimental conditions were dummy-coded independent variables, each of them contrasting with the non-treatment condition; all outcome measures were statistically adjusted for History Relevance pretests and six students' background variables as covariates: educational level, grade, age, gender, history outside class and historical knowledge. 'History outside class' refers to the extent to which students talked about history at home or with other people, which was measured by one item with a 4-point Likert scale ranging from 'never' to 'often'. 'Historical knowledge' refers to general historical knowledge test students took prior to the intervention.

An a priori power analysis showed adequate statistical power to test our hypotheses ($\beta > .80$) for small, medium and large effects, according to the rules of thumb of Cohen (1988). All tests were conducted at the conventional alpha level of .05. As we assumed no differences regarding the fourth hypothesis, the corresponding test was not aimed at rejecting the null hypothesis, and we therefore adjusted the alpha level to $\alpha = .50$.

Model fit was evaluated with the log-likelihood test (differences between $-2LL$ of the intercept-only model and the final model) and ‘explained’ variance. Effect sizes (Cohen’s d) were calculated for statistically significant effects related to our main hypotheses. Standardized model-based effect sizes were calculated, expressing experimental differences after adjustment for the covariates from our statistical models. We also determined effect sizes based on observed scores (d_{obs}), expressing the experimental differences for the observed outcome measures.

5.9 Results

5.9.1 Descriptive statistics

Table 5.4 shows the descriptive statistics. Regarding the general historical knowledge test, mean score differences between groups were small and statistically insignificant, as mentioned earlier. Pretest mean scores for History Relevance-total were above average for all three groups, which can be attributed to the relatively high scores for Relevance-citizenship and Relevance-human condition. Situational Interest-total mean score was higher for the case-comparison group than for the separate-case group, which was mainly due to the relatively high scores for Maintained-SI-value and Triggered-SI. The case-comparison group had experienced less difficulty with the applied approach than the separate-case group as indicated by their Pedagogical Approach scores. Both groups correctly answered about 73 percent of the Lesson Content Knowledge items.

5.9.2 Students’ appraisals of the relevance of history

Results of multilevel analyses conducted for History Relevance are presented in Table 5.5. The case-comparison condition showed significant effects for Relevance (total), Identity, Citizenship and Human Condition with small but significant model-based effect sizes of 0.09, 0.08, 0.08 and 0.06 respectively ($d_{\text{obs}} = 0.17, 0.13, 0.17$ and 0.12). The separate-case condition showed significant effects for Relevance (total) and Citizenship with significant model-based effect sizes of 0.06 and 0.07 respectively ($d_{\text{obs}} = 0.15$ and 0.21). Acknowledging the small sizes of the effects, these outcomes are in line with our hypothesis that students’ appraisals of the relevance of history were positively affected to a greater extent in the case-comparison condition than in the separate-case condition.

Table 5.4 Mean scores and (standard deviations) for the measures for each group.

	N	<i>Case-comparis on group</i>	N	<i>Separate- case group</i>	N	<i>Non- treatment group</i>
Historical Knowledge	460	25.61 (5.03)	273	25.49 (4.47)	289	25.11 (4.73)
History Relevance: <i>pretest</i>						
Total	460	3.78 (0.70)	273	3.79 (0.70)	289	3.79 (0.72)
Identity	460	3.03 (0.90)	273	2.98 (0.88)	289	3.08 (0.92)
Citizenship	460	4.06 (0.74)	273	4.09 (0.74)	289	4.08 (0.76)
Human condition	460	4.14 (0.85)	273	4.18 (0.87)	289	4.13 (0.91)
History Relevance: <i>posttest</i>						
Total	460	3.85 (0.76)	273	3.83 (0.70)	289	3.72 (0.76)
Identity	460	3.22 (0.94)	273	3.15 (0.87)	289	3.10 (0.94)
Citizenship	460	4.07 (0.81)	273	4.09 (0.74)	289	3.93 (0.80)
Human condition	460	4.18 (0.84)	273	4.15 (0.87)	289	4.08 (0.89)
Situational Interest (SI)						
SI (total)	444	3.21 (0.76)	264	3.12 (0.66)	-	-
Maintained-SI-Value	444	3.37 (0.76)	264	3.24 (0.72)	-	-
Maintained-SI-Feeling	444	3.17 (0.91)	264	3.12 (0.82)	-	-
Triggered-SI	444	3.10 (0.83)	264	3.00 (0.71)	-	-
Pedagogical approach	444	2.50 (0.78)	264	2.75 (0.77)	-	-
Lesson Content Knowledge	444	22.18 (3.60)	264	21.71 (4.27)	-	-

Table 5.5 Multilevel models for History Relevance ($N = 1022$).

	Relevance (total)	Identity	Citizenship	Human Condition
Fixed model				
Intercept	3.751 (.037)*	3.064 (.060)*	4.005 (.043)*	4.083 (.053)*
Case-comparison condition ^a	.130 (.032)*	.121 (.052)*	.125 (.037)*	.097 (.046)*
Separate-case condition ^a	.104 (.035)*	.068 (.057)	.117 (.041)*	.038 (.051)
<i>Covariates</i>				
Educational level ^b	-.028 (.029)	-.078 (.046)	-.006 (.033)	.038 (.041)
Grade ^c	.016 (.035)	.042 (.057)	-.011 (.041)	.042 (.050)
Age (centered)	.023 (.016)	.036 (.026)	.025 (.019)	.011 (.023)
Gender ^d	-.058 (.027)*	.051 (.044)	-.109 (.031)*	-.057 (.039)
History outside class	.046 (.021)*	.100 (.033)*	.099 (.024)*	.080 (.028)*
History knowledge	-.001 (.003)	.001 (.005)	.000 (.004)	.000 (.005)
Corresponding relevance pretest	.855 (.023)*	.732 (.027)*	.778 (.024)*	.691 (.023)*
Random model				
	Variance (SE)	Variance (SE)	Variance (SE)	Variance (SE)
Class level	.040 (.016)* - .000 (.000)	.034 (.017)* - .000 (.000)	.053 (.020)* - .000 (.000)	.031 (.014)* - .000 (.000)
Student level	.520 (.023)* - .168 (.007)*	.904 (.040)* - .440 (.019)*	.583 (.026)* - .227 (.010)*	.713 (.032)* - .344 (.015)*
R^2	.70	.53	.64	.54
-2 LL for null model and final model and Δ - 2LL ($\Delta df = 9$)	2264.420 – 1075.946 (1188.474*)	2818.950 – 2060.851 (758.099*)	2384.131 – 1385.881 (998.250*)	2577.920 – 1809.776 (768.144*)

$$R^2 = (\sigma_{\text{null model}} - \sigma_{\text{estimated model}}) / \sigma_{\text{null model}}$$

* $\rho < .05$

a Compared to the non-treatment group; b Reference category = pre-university level; c Reference category = grade 10; d Reference category = male

As expected, History Relevance pretest scores proved to be robust predictors of History Relevance posttest scores. There were no significant correlations between History Relevance posttest scores and educational level, grade, age and historical knowledge. Male students found history less relevant than female students. History outside class significantly predicted all History Relevance outcomes, implying that the extent to which students talk about history in extracurricular contexts was positively related to their views on the relevance of history.

Log-likelihood tests showed significant fit for all four History Relevance outcomes measures. The explained variance ranged from $R^2 = .53$ to $.70$, indicating strong explanatory power of the models.

5.9.3 Students' situational interest (SI)

Multilevel models for SI are shown in Table 5.6. The separate-case condition was not significantly related to any of the SI outcomes whereas the case-comparison condition showed a small but significant effect on Maintained-SI-Value (model-based effect size: 0.08 , $d_{\text{obs}} = 0.18$). This means that the situational interest of both groups did not differ in terms of attention, engagement and enjoyment. However, the case-comparison group deemed the lessons more valuable than the separate-case group.

History Relevance pretest substantially predicted SI posttest-scores, indicating a strong positive correlation between students' views on the relevance of history and their interest regarding the lesson unit. No significant relationships were found for educational level, gender and history outside class. Grade positively correlated with Maintained-SI-Feelings, implying that grade 10 students considered the course materials more enjoyable and engaging than grade 11 and 12 students. Historical knowledge negatively related to Triggered-SI, implying that the lesson unit grabbed less attention from students with a lower knowledge level.

Log-likelihood tests showed significant fit for all four SI-models with explained variance ranging from $R^2 = .11$ to $.21$.

Table 5.6 Multilevel models for Situational Interest ($N = 708$).

	Situational Interest (SI) (scale)	Maintained-SI-Value (subscale)	Maintained-SI-feeling (subscale)	Triggered-SI (subscale)
<i>Fixed model</i>				
Intercept	3.026 (.080)*	3.170 (.081)*	3.007 (.096)*	2.903 (.084)*
Case-comparison condition ^a	.117 (.074)	.159 (.075)*	.091 (.088)	.094 (.076)
<i>Covariates</i>				
Educational level ^b	.081 (.079)	.067 (.080)	.094 (.094)	.088 (.083)
Grade ^c	.152 (.083)	.121 (.083)	.239 (.100)*	.108 (.089)
Age (centered)	.046 (.032)	.014 (.032)	.055 (.039)	.073 (.036)*
Gender ^d	-.054 (.052)	-.066 (.053)	-.120 (.065)	.019 (.059)
History outside class	.033 (.042)	.023 (.043)	-.048 (.052)	-.028 (.048)
History knowledge	-.009 (.006)	-.001 (.006)	-.009 (.008)	-.017 (.007)*
Relevance (total) pretest	.441 (.045)*	.505 (.045)*	.454 (.055)*	.346 (.050)*
<i>Random model</i>				
Class level	Variance (SE) .042 (.017)* -.026 (.012)*	Variance (SE) .043 (.017)* -.027 (.012)*	Variance (SE) .049 (.021)* - .033 (.016)*	Variance (SE) .040 (.017)* -.021 (.012)
Student level	.491 (.027)* - 417 (.023)*	.527 (.029)* - 422 (.023)*	.723 (.039)* - .644 (.035)*	.589 (.032)* - .541 (.029)*
R^2	.17	.21	.12	.11
-2 LL for null model and final model and Δ	1538.526 - 1416.796	1587.687 - 1425.544	1807.939 - 1720.460	1662.643 - 1592.170
-2LL ($\Delta df = 8$)	(121.730*)	(162.143*)	(87.479*)	(70.473*)

$$R^2 = (\sigma_{\text{null model}} - \sigma_{\text{estimated model}}) / \sigma_{\text{null model}}$$

* $\rho < .05$

a Compared to the separate-case condition; b Reference category = pre-university level; c Reference category = grade 10; d Reference category = male

5.9.4 Students' experiences with the pedagogical approach (PA)

In line with hypothesis 3, the case-comparison condition had a significant negative effect on PA outcome, implying that students in this condition considered the applied pedagogical approach less difficult than students in the separate-case condition (see Table 5.7); model-based effect size was 0.15 ($d_{\text{obs}} = 0.32$, corresponding to a small-to-medium effect). There were also significant correlations for gender (male students experiencing less difficulty with the approach than female students), historical knowledge and History Relevance pretest scores, whereas educational level, grade, age and history outside class were not significantly related to PA. A log-likelihood test showed adequate model fit with an explained variance of $R^2 = .14$.

5.9.5 Acquiring lesson content knowledge (LCK)

In line with hypothesis 4, there were no meaningful differences between the case-comparison and separate-case group in terms of knowledge acquisition (see Table 5.7) as condition had no significant effect on LCK ($t = -0.250, p = .422$). Educational level, grade, gender, historical knowledge and History Relevance pretest scores were significant covariates, meaning, for example, that pre-university students (VWO) performed better than middle level general secondary students (HAVO) and students with a relatively high level of general historical knowledge performed better than students with a relatively low knowledge level. Male students outperformed female students and grade 10 students underperformed grade 11 and 12 students. A log-likelihood test showed adequate model fit with an explained variance of $R^2 = .21$.

Table 5.7 Multilevel models for Pedagogical Approach and Lesson Content Knowledge ($N = 708$).

	Pedagogical Approach	Lesson Content Knowledge
<i>Fixed model</i>		
Intercept	2.916 (.084)*	20.445 (.486)*
Case-comparison condition ^a	-.219 (.077)*	-.115 (.459)
<i>Covariates</i>		
Educational level ^b	-.139 (.083)	1.956 (.478)*
Grade ^c	-.121 (.088)	-1.152 (.467)*
Age (centered)	.019 (.035)	-.119 (.159)
Gender ^d	-.165 (.058)*	.575 (.261)*
History outside class	-.037 (.047)	.232 (.210)
History knowledge	-.023 (.007)*	.205 (.032)*
Relevance (total) pretest	-.193 (.049)*	.788 (.223)*
<i>Random model</i>		
	Variance (SE)	Variance (SE)
Class level	.064 (.023)* - -.023 (.012)	3.016 (.915)* - 1.577 (.529)*
Student level	.555 (.030)* - 511 (.028)*	11.942 (.651)* - 10.177 (.555)*
R^2	.14	.21
-2 LL for null model and final model and Δ	1632.108 - 1556.167	3818.513 - 3692.769
-2LL ($\Delta df = 8$)	(75.941*)	(125.744*)

$$R^2 = (\sigma_{\text{null model}} - \sigma_{\text{estimated model}}) / \sigma_{\text{null model}}$$

* $\rho < .05$

a Compared to the separate-case condition; b Reference category = pre-university level; c Reference category = grade 10; d Reference category = male

5.10 Conclusion and discussion

In this study, we examined the efficacy of history teaching about an enduring human issue by means of comparing analogous cases from different periods. In both experimental conditions, this way of teaching had a positive effect on students' appraisals of the relevance of history compared to the relevance perceptions of students who followed the regular history curriculum. As we hypothesized, students' relevance appraisals were positively affected to a greater extent in the case-comparison condition than in the separate-case condition. The case-comparison group also considered the lesson unit more valuable and had less difficulty with the applied pedagogical approach, which contrasted strongly with the history curriculum students were accustomed to. They did not underperform in terms of knowledge acquisition, even though the focus was on comparison activities and drawing analogies between past and present whereas learning activities in the separate-case condition aimed at gaining historical knowledge and practicing historical thinking skills. These results are in line with cognitive psychology research literature evincing positive effects of case-based comparison activities and meaningful contexts on learning performances and subject matter value perceptions (e.g. Alfieri et al., 2013; Gentner, Loewestein, & Thompson, 2003).

Overall, students in both treatment groups as well as in the non-treatment group associated the relevance of history with 'becoming a citizen' and 'understanding the human condition' rather than with 'building a personal identity'. These results comply with findings from a large-scale European survey on students' attitudes towards historical consciousness (Angvik & Borries, 1997), revealing that, on average, students see more relevance in history for orienting on the present and future than for guiding their individual lives. The results also comply with a Dutch 'school history experience' survey conducted among grade 11 and 12 middle level general secondary and pre-university secondary students, according to which a large majority believed that history is irrelevant for learning more about the self (Van der Kaap & Folmer, 2016). Given the mean age of participants (about 16 years), our findings are also in line with claims that identity building processes occur late in adolescence and sometimes not even until young adulthood (e.g. Steinberg & Morris, 2001). However, it has been demonstrated in this study that students' attitudes towards the relevance of history in terms of building a

personal identity can be influenced, even when controlling for students' background characteristics.

Although the intervention yielded statistically significant effects on students' appraisals of the relevance of history, effect sizes were small. There are some obvious explanations. Consistent with other empirical findings (Van Straaten, Wilschut, & Oostdam, 2018; Van der Kaap & Folmer, 2016), students in the treatment groups as well as in the non-treatment group deemed history considerably relevant at the start of the experiment, leaving limited room for improvement; the fact that students had opted for a curriculum including history has probably been of influence. Although the intervention was spread over several weeks and involved pedagogical approaches students were unfamiliar with, its duration was actually relatively short, reducing the chance of generating impact and change in the classroom (Slavin, 2008). Furthermore, measuring interventional effects on students' history relevance appraisals presupposes a change of opinions and attitudes. Attitude change depends on a complex interplay of multiple factors, such as a person's need for social acceptance, motivation, the complexity of knowledge underlying existing beliefs or the ability for processing information that might influence those beliefs (Fabrigar, Petty, Smith, & Crites, 2006; Mason, 2001; Petty & Wegener, 1998; Wood, 2000). As both cognitive and affective factors play a role here, focusing on cognitive learning activities - like in this study - does not guarantee attitude modification. Moreover, research has indicated greater stability in attitudes for older than younger individuals (e.g. Alwin & Krosnick, 1991). Given all of these considerations, there is reason to believe that the small increases in relevance appraisals detected in this study may be more meaningful than their modest size suggests. Also, it would be worthwhile to further experiment with this study's pedagogical approaches in junior secondary education, not only because junior secondary students are presumably more susceptible to attitude change, but also because studies have shown that junior students have lower history relevance appraisals than their senior peers (Haeberli, 2005; Haydn & Harris, 2010; Van Straaten, Wilschut, & Oostdam, 2018; Wilschut, 2013).

Three limitations of this study should be borne in mind. First, not all classes were randomly assigned to the experimental conditions and operating in a natural educational setting may have been influenced by potential differences between groups. We have reduced selection bias by (1) using multilevel analysis controlling for several background

characteristics, (2) determining equivalence between all groups at the pretest, and (3) using a relatively large sample size (Shadish, Cook, & Campbell, 2002; Slavin, 2008).

Second, although students in the separate-case group were not actively encouraged to compare cases, it cannot be ruled out that they did make comparisons, even though in educational contexts people are not inclined to compare analogous situations spontaneously (Gentner, Loewenstein, & Thompson, 2003; Rittle-Johnson & Star, 2007). Two out of 11 teachers participating in the separate-case condition reported that students occasionally referred to current refugee issues, but, according to instructions, these teachers did not respond and summoned the students to focus on lesson contents.

Third, the topic of the lesson unit - migration and integration – referred to current affairs in the Netherlands during the lesson intervention due to Syrian refugees seeking asylum in Europe. This may have affected the outcomes – either in a positive way by triggering students’ engagement or in a negative way by evoking feelings of aversion or saturation. Therefore, further research should deploy various enduring issues – including less morally laden – to test the generalizability of the observed effects in this study.

If history is to contribute to the education of students on their path to citizenship, knowledge of the past should be connected to perceptions of the present and expectations for the future. To date, there is a lack of empirically-based pedagogical tools teachers can use for making such connections. This study was undertaken to fill this gap by examining the efficacy of analogical reasoning and case comparison learning in the context of teaching an enduring human issue. The findings show that these pedagogical approaches are powerful tools to enhance the quality of history teaching. Comparing ways in which people in the past responded to enduring issues provides students with opportunities to develop thoughts about contemporary dealings with similar issues, thus using historical knowledge for their orientation on present and future. These are important yields for practitioners who are pursuing meaningful history teaching and trying to enhance student motivation.