Toward a framework for assessing the quality of students' social scientific reasoning

Klijnstra, T.; Stoel, G.L.; Ruijs, G.J.F.; Savenije, G.M.; van Boxtel, C.A.M.

DOI
10.1080/00933104.2022.2132894

Publication date
2023

Document Version
Final published version

Published in
Theory and Research in Social Education

THEORY & RESEARCH IN SOCIAL EDUCATION
2023, VOL. 51, NO. 2, 173–200
https://doi.org/10.1080/00933104.2022.2132894

Routledge
Taylor & Francis Group

Toward a framework for assessing the quality of students' social scientific reasoning

Thomas Klijnstra[a], Gerhard L. Stoel[b], Gerard J. F. Ruijs[a], Geerte M. Savenije[a], and Carla A. M. van Boxtel[a]

[a]University of Amsterdam; [b]Radboud University Nijmegen

ABSTRACT

This study aims to describe components and levels of upper secondary social science students' reasoning about social problems. We consulted conceptualizations of social scientific reasoning in sociology textbooks and social science education literature, analyzed student papers, and conducted focus groups with social science teachers and teacher educators to define social scientific reasoning by proficiency levels and identify common flaws in students' reasoning. The papers were written by upper secondary social science students from eight schools in the Netherlands. We defined social scientific reasoning in terms of three components (describing, explaining, and solving problems) and five reasoning activities (causal analysis; use of social scientific concepts, models, and theories; use of evidence; use of perspectives and reflections on them; and comparing). We described these reasoning activities in three proficiency levels supported by practical examples and rubrics for students' reasoning. These insights can inform teachers and teacher educators in monitoring students' progression and designing teaching materials and activities that can promote students' social scientific reasoning.

KEYWORDS

Citizenship education; flaws in reasoning; reasoning levels; social science education; social scientific reasoning; social studies education

> At a broader societal level, a democracy composed of citizens who can think for themselves on the basis of evidence and concomitant analysis, rather than the emotion, prejudice, or dogma, is a plus—in fact, it sustains, builds, and perpetuates the democracy. (Abrami et al., 2008, p. 1103)

Social science education is critical to help students make sense of society. Although the name and content of this subject may vary from country to country, the familiar principle in social science education involves teaching students to understand and analyze current social problems (e.g., refugee crises, growing inequality) and employ reasoning to arrive at potential solutions (Barton & Avery, 2016; Barton, 2011; Sandahl, 2015). This study is conducted in the Netherlands. Sociological concepts form the basis of the Dutch social science school subject "Maatschappijwetenschappen." This school subject teaches students to use social science concepts and theories to analyze current social problems and think about and discuss possible solutions (Klijnstra et al., 2022). For example, students learn to analyze the relationship between people's use of media and the use of violence in a particular context. More specifically, students should relate different causes and

CONTACT Thomas Klijnstra t.klijnstra@uva.nl Research Institute of Child Development and Education (RICDE), University of Amsterdam, PO Box 15776, Amsterdam 1001 NG, The Netherlands.

consequences to each other and substantiate their arguments with empirical evidence. In sum, students need to *reason* about social problems in a social scientific way.

The more general literature has shown reasoning to involve making claims, providing evidence, and analyzing the logic in between (e.g., Cavagnetto & Hand, 2011; Toulmin, 1958). However, reasoning is often challenging; previous research showed that, in general, humans reason erratically, make apparent errors, and are biased in decision making (e.g., Evans, 2002; Lee et al., 2021). People make firm statements about social and political issues but often can not substantiate their opinions. People often search for evidence that fits their beliefs (e.g., Fernbach et al., 2013), and their reasoning is usually based on gut feelings and misunderstandings (or oversimplifications) of socio-political issues (Dawson & Venville, 2009; Fernbach et al., 2013; Kuhn, 2007).

In different school subjects, students learn to reason scientifically about certain contents. For example, students reason about natural phenomena in physics or about social phenomena in social studies. The term "scientific reasoning" can be used as an umbrella term for reasoning in different disciplines. Characteristics of scientific reasoning are learning to define a problem, formulate a hypothesis, evaluate evidence, and formulate a conclusion. This process can be worked out more specifically for each school subject. What it means to reason scientifically in a domain specific way has been further conceptualized for science, technology, engineering, and mathematics (see Bao et al., 2009; Sadler, 2004) and historical education (see Seixas & Morton, 2012; Van Boxtel & van Drie, 2018; Wineburg, 2001). However, for social sciences education this conceptualization of reasoning is lacking.

In the United States, the National Academy of Education published a report on civic reasoning with the aim to better prepare students to examine and discuss complex civic, political, and social issues. This civic reasoning focuses on public issues in which students must use inquiry skills, select and use evidence, and weigh multiple perspectives with the central question: what should students do? (Lee et al., 2021). The term civic reasoning is thus used to refer to reasoning about social or socio-scientific issues, which can be practiced in various subjects, including, for example, biology, geography, social sciences, and history. Therefore, Lee et al. (2021) recommended that further in-depth research is needed on reasoning in specific academic disciplines, which is also in line with previous research on the relevance of subject-specific components of reasoning (see Goldman et al., 2016) and the *College, Career, and Civic Life (C3) Framework for Social Studies State Standards* (National Council for the Social Studies [NCSS], 2013). In the *C3 Framework*, applying disciplinary knowledge and tools is considered part of conducting research in social studies.

This study is a subject-specific elaboration of both scientific reasoning and civic reasoning, namely, within the field of social science education. This study examines students' reasoning about current social issues in social science education. We labeled this reasoning social scientific reasoning to indicate both the social scientific approach (in the Netherlands, using primary concepts and methods from sociology) and the societal content and focus of this type of reasoning.

Empirical evidence on upper secondary school students' social scientific reasoning is scarce, and little is known about the difficulties students encounter when analyzing social problems. In this study, we aim to conceptualize social scientific reasoning in terms of concrete components, activities, and levels, using both the literature and the reasoning that students demonstrate in written papers. We evaluate the usefulness of this conceptualization by examining whether it can be used to assess the quality of students' reasoning in the

context of a specific assignment. The components and activities, and how they may differ in quality, are described in a rubric with which the quality of students' reasoning can be assessed. Insights from this study contribute to the empirical basis of a definition of social scientific reasoning and enable research on progress and intervention studies. Finally, these insights are relevant for social science teachers and teacher educators for fostering—and assessing the quality of—students' reasoning, with the hope that improved social scientific reasoning will lead to better-equipped future citizens.

Theoretical framework

The function of reasoning is generally seen to be increasing knowledge and making better decisions. For high-quality reasoning, domain-specific knowledge and skills are required (Fischer et al., 2014; Goldman et al., 2016; Lee et al., 2021; NCSS, 2013). In social science education, students reason about contemporary social problems. Mills (2000) defined social problems as issues involving many people, transcending the individual. Often, there are different views on the problem's solution due to its complexity and to conflicting values and interests (Mills, 2000; Ultee et al., 2003; van Tubergen, 2020). However, while this definition seems straightforward, identifying a problem as a social one is highly dependent on context and time (van Tubergen, 2020). Crime is, for example, a social problem; it transcends the individual and affects many people, and there are different views on approaches (i.e., solutions) to crime. Yet crime is heavily dependent on both time and context. For example, adultery is a criminal offense in some countries but not in others.

In sociology, the focus is also on analyzing social problems. Within sociology, there is a threefold division regarding the study of social problems, namely: to *describe* social problems in order to understand them more accurately, to *explain* them using theories and models, and to *apply* sociological insights to make predictions and/or to develop and evaluate social interventions (Ultee et al., 2003; van Tubergen, 2020). As van Tubergen (2020) has stated, "Generally speaking, one can say that the aim of sociology is to come up with accurate scientific descriptions and theoretical explanations for social phenomena, and to apply their knowledge" (p. 16).

This threefold division within the field of sociology also plays a vital role in the literature for analyzing social problems in social science education. Sandahl (2015) presented a model "analyzing societal issues using causes and consequences" that contained four phases that students must enact when analyzing societal issues: *describing the problem* of a situation, *describing the causes* underlying the problem, *describing the consequences* of the problem, and *describing the solutions* to the problem. The third phase is closely related to the first. When problems are described, their consequences must also be specified.

The identification of the main components of reasoning about social problems needs to be further elaborated in terms of reasoning activities. What, for example, are students actually doing when they explain social problems? Based upon interviews with six experienced social science teachers, and applying a theoretical framework on second-order concepts in history, Sandahl (2015) identified six second-order thinking concepts and related activities: analyzing cause and consequence, using evidence and making inference, using abstractions (models and theories) to understand, comparing and contrasting, taking perspectives, and analyzing the evaluative dimension.

Building on the conceptualizations of van Tubergen (2020), Ultee et al. (2003), and Sandahl (2015), we conceptualized students' social scientific reasoning in terms of three main components: describing, explaining, and solving social problems. Solving the problem does not refer to actual resolution; instead, students learn to weigh different points of view and discuss possible solutions. Based on the conceptualizations of domain-specific aspects of reasoning (Goldman et al., 2016; Lee et al., 2021) and subject-specific reasoning (Sandahl, 2015; Ultee et al., 2003; van Tubergen, 2020), we identified five distinct reasoning activities that students can engage in on different levels: causal analysis; use of social scientific concepts, models, and theories; use of evidence; use of perspectives and reflection on them; and comparing.

At first glance, these five reasoning activities are well-known categories. For example, previous research on historical reasoning identifies six components of students' historical reasoning: asking historical questions, using sources, contextualizing, using argumentation, using substantive concepts, and using meta-concepts (van Drie & Van Boxtel, 2007). Research on civic education reveals similar but distinctive elements of civic reasoning. The emphasis in civic reasoning is on current issues and on the roles that students can play in those issues. Civic reasoning, for example, involves individuals' reasoning on public matters to create consensus and mutual understanding (Lee et al., 2021). However, the five social scientific reasoning activities we identified above differ from these historical or civic reasoning activities because of their different aims and contents. We will now illustrate these reasoning activities with examples and highlight similarities and differences compared to other school subjects and to previous research.

Causal analysis

Causal analysis has generic but also subject-specific features. In history, causal analysis is, among other things, about learning to reason with multiple causes and effects (Seixas & Morton, 2012). In contrast, in science education, causal analysis is primarily about isolating one variable. Research showed that students find it difficult to reason with more than one cause; they often reason too linearly, and they may overestimate the role of human action in causal analysis (Coffin, 2004; Seixas & Morton, 2012; Stoel et al., 2015).

Such insights (reasoning with multiple causes, the role of human action, and learning to reason less linearly) are also recognized when analyzing social problems (Sandahl, 2015; van Tubergen, 2020). However, social scientific reasoning comes with additional challenges; for example, students must also be able to link insights about causes and effects to possible solutions. In addition, students need to learn that various social science theories and concepts also include causal analysis. Lundholm and Davies (2013) have illustrated that the use of social science theories are seldom about direct causality. When using concepts, more attention needs to be paid to causality and context in the social sciences. As far as our knowledge goes, insights into levels of social scientific reasoning about causes and consequences of social problems are lacking.

Use of evidence

The effective use of evidence is a critical component of reasoning. After all, statements must be substantiated with proof (Goldman et al., 2016). However, the use of evidence can be

challenging. Students find it difficult to link data to claims when reasoning in the context of science education (Jiménez-Aleixandre et al., 2000; Kelly et al., 1998; McNeill et al., 2006). Furthermore, studies on the use of evidence in history education have shown that students often provide few arguments for their claims (McCarthy Young & Leinhardt, 1998; Monte-Sano, 2010), and students have great difficulty analyzing and integrating evidence into their texts (De La Paz & Felton, 2010; Monte-Sano, 2010; Stoel, 2017; Wineburg, 1991). Recent research on civic online reasoning has shown that students find it difficult to evaluate online sources and to distinguish between advertisements and news items (McGrew et al., 2018; Wineburg & McGrew, 2016).

Due to the object of empirical inquiry in social sciences and the focus on human behavior, there is often a greater level of uncertainty in the evidence compared to the hard sciences (Acar et al., 2010). Students need to use statistics and interviews to obtain evidence about social problems in which the quality of measurement, validity, and reliability determine how convincing the evidence can be (Avery & Barton, 2017; van Tubergen, 2020). This task is far from straightforward. Students routinely have encountered problems evaluating social scientific evidence (Acar et al., 2010; Sadler, 2004).

Most research on the use of evidence has focused on describing or explaining social problems. However, we also need to define how students use evidence to reason about the effectiveness of possible solutions concerning the problem at hand. Nevertheless, empirical insights on this subject are still lacking.

Use of social scientific concepts, models, and theories

When analyzing social problems, it is essential to use concepts, models, and theories developed in the social sciences to understand the problem (Sandahl, 2015; van Tubergen, 2020). This application of concepts, models, and theories has also been emphasized in the issues centered approach in which students appropriate subject-specific knowledge and skills in the context of real issues that expose relevant social dilemmas (e.g., Evans et al., 1996). The National Academy of Education has advised that students should understand and apply subject specific concepts and knowledge to historical, political, social, and economic issues, among others (Lee et al., 2021).

Conceptual models can uncover the relationships between various concepts (Jacoby & Jaccard, 2010; van Tubergen, 2020). However, these abstractions refer to specific instances (van Tubergen, 2020). For example, social inequality could refer to a country's income inequality or to the affordability of healthcare. Students should learn to connect abstract sociological and political science concepts to specific social contexts and be capable of concretizing these concepts. For example, such concepts become more concrete when students learn to distinguish appropriate indicators.

Several scholars have shown that content knowledge affects the quality of reasoning (Flemming, 1986; Sadler & Zeidler, 2004; Tytler et al., 2001), with more content knowledge leading to a higher level of reasoning (Hogan, 2002; Sadler & Zeidler, 2004). Moreover, research has shown that young adults find it challenging to use theories in their reasoning that contradict their own beliefs (Chinn & Brewer, 1993; Kuhn, 1991). The use of social scientific concepts, models, and theories are distinct from reasoning in other school subjects. For example, the use of models and theories in history education is scarce. Yet,

empirical evidence of students' use of social scientific concepts, models, and theories is lacking.

Use of perspectives and reflection on them

There can be different perspectives on social problems and how they can be explained and resolved (van Tubergen, 2020). A higher level of reasoning is related to comprehensively understanding the context of a problem. Problems can be more accurately understood by learning to examine them from political-legal, socio-cultural, and socio-economic perspectives (Sandahl, 2015, 2020). Thus, evaluating normative and conflicting goals plays an important role when students analyze social problems.

When students reason about social problems, their emotions can affect their reasoning (Sitzlein, 2021), which is consistent with previous research among adults; reasoning can be influenced by people's views and by personal engagement (Kunda, 1990; Mercier & Sperber, 2011). When reasoning, people compare and contrast their own opinions with those of others and judge the quality of arguments that confirm their point of view to be better than those that contradict it. This phenomenon is called *motivated reasoning* (Kunda, 1990). Research suggests that students are apt to fall victim to motivating reasoning when analyzing data; for example, Zummo et al. (2020) provided evidence for ideologically motivated reasoning among youth about climate change (Zummo et al., 2020).

Similarly, when people give their opinions on social problems, they seek out information that fit their points of view, avoiding information and arguments contrary to their own opinions and perspectives (Kunda, 1990; Taber & Lodge, 2006). People are biased information processors. For example, when participants read pro and con arguments about gun control, they counter the opposing arguments and uncritically accept the supporting arguments—evidence of a confirmation bias (Taber & Lodge, 2006). Moreover, people interpret contradictory information to match their points of view and tend to confirm their position when reasoning about specific issues (Kunda, 1990; Taber & Lodge, 2006; Zummo et al., 2020). Research in learning about controversial issues in history education also showed that when students are personally more involved in a specific topic, they find it more challenging to engage with other perspectives (Barton & McCully, 2012; Goldberg & Savenije, 2018; Goldberg, 2013). In social science education, students should learn to reason with distance, compare perspectives, and be able to reflect on their involvement (Sandahl, 2020), which may be especially challenging and relevant when it comes to reasoning about current issues because students may hold preexisting views, opinions, and emotions.

Comparing

Comparing is a central component of social science thinking and reasoning (van Tubergen, 2020; Wilterdink et al., 2017). As Sandahl (2015) illustrated, comparing (and contrasting) can help students understand phenomena better. For example, by asking students to compare social inequality in the context of Dutch and U.S. high schools, they can better understand the factors that can affect the degree of inequality in a society. To achieve a higher level of comparison, students must learn to categorize (e.g., operationalize indicators for comparison) and describe similarities and differences between certain contexts (Ruijs & Klijnstra, 2021; van Tubergen, 2020). A scientific attitude in general, and

specifically a social scientific attitude, requires students to question conclusions by comparing and contrasting them with other possible conclusions (Ruijs & Klijnstra, 2021; van Tubergen, 2020).

Research questions

This study aims to conceptualize students' social scientific reasoning in terms of activities as concrete as possible and validated in student work. A conceptualization of social scientific reasoning in terms of concrete behavior related to describing, explaining, and solving social problems, and the development of performance indicators, should make it possible to assess the quality of students' social scientific reasoning in the classroom. Therefore, we developed our conceptualization in the form of rubrics and investigated whether—using these rubrics—we can assess the quality of students' reasoning in the context of a concrete assignment. Accordingly, we developed the following research questions: How can we define students' social scientific reasoning activities in terms of concrete behavior while distinguishing between different levels of quality? Can this conceptualization be used to assess the quality of students' reasoning in the context of a specific assignment?

Method

Context of the study

We conducted the study on Dutch social science education. In the Netherlands, social science is an exam subject in upper secondary school. The subject can be taken by those in the two highest educational tracks: higher general continued education (those preparing for a university of applied sciences) and pre-university education (those preparing for a research university). In the Netherlands, pre-university education takes six years to complete (7th to 12th grade) while higher general education takes five years to complete (7th to 11th grade).

In 2017, a new social science curriculum—the primary innovation of which was applying a concept-context approach—was introduced in the Netherlands. The starting point of the new curriculum is that students learn to apply their knowledge of social scientific concepts in meaningful contexts. For example, students must learn to analyze current social issues using socio-political concepts. The new examination program emphasizes the importance of disciplinary knowledge and critical thinking about real-life problems (College van Toetsen en Examens, 2019). This revised program emphasizes first-order concepts (e.g., globalization and social inequality) and second-order concepts (e.g., causation and evidence). In contrast to the previous examination program, second-order concepts have been afforded greater importance.

Data collection and participants

Using the three components and five reasoning activities we described in the theoretical framework, we developed a rubric for each component that includes the five reasoning activities elaborated in concrete indicators and levels. The data collection and analysis consisted of four different phases: exploration phase, construction phase, validation and

reconstruction phase, and evaluation phase (McKenney & Reeves, 2018). In these four phases, we randomly selected and analyzed a total of 83 student papers from a complete set of 388 papers collected during a previous study (Van Boxtel et al., 2017). These papers were written by Dutch students in upper secondary social science education (with a mean age of 16). These students originated from eight schools in different areas of the Netherlands, and there were eight assignments in which students were challenged to reason about social problems. The assignments were developed in a teacher development program (TDP) in the context of the previously mentioned curriculum revision for the subject of social sciences. In the TPD teachers were guided by two social science teacher educators and two researchers. The question in this TDP was how to adequately teach and assess students' reasoning about social problems by the use of authentic learning assignments, which are meaningful, realistic (larger) assignments that require more complex skills from students (Maddox & Saye, 2017). Therefore, analyzing these assignments was of interest to this current study of the quality of social science reasoning. In five of the eight assignments, students worked individually. In three of the eight, they worked in groups (two, three, or four students). Students in the pre-university track were, on average, one year older than those in the senior general secondary education track. Moreover, the pre-university students had received an additional year of social science education, resulting in an average extra study load of 140 hours. The student papers were written in the context of different authentic assessment assignments about social problems. In the papers (each of which was one to three pages in length), the students were asked to, for example, give advice on the accommodation of refugees in their city. In another assignment, students compared the conclusions drawn by Wilkinson and Pickett (2010) about the relationship between social inequality and social problems with empirical data from other researchers on the same social problems, such as life expectancy, homicide, the degree of trust in a society, and obesity levels.

Exploration phase

We conducted a literature review on sociology and social science education to define social scientific reasoning activities. The literature review enabled us to derive insights into the difficulties of, and flaws in, reasoning. We read and discussed 30 (randomly selected) student papers in several cycles using the reasoning activities from our model as sensitizing concepts. We analyzed the student papers (four to five papers per cycle), highlighting examples of reasoning we considered *weak*, *average*, or *strong*, and collected examples of frequently-occurring flaws.

Construction phase

Based on the analysis of the first set of 30 student papers, we constructed initial rubrics and a list of reasoning flaws. We ensured that the descriptors for each level had the same performance criteria and attributes for the progressive scale to be continuous and consistent between levels (Simon & Forgette-Giroux, 2001). Based on this initial set, we refined our classification in subcategories of students' social scientific reasoning and tested it on the second set of 30 papers. Accordingly, this process greatly informed us about the different categories and levels of students' reasoning about social problems. The analysis of this first set of student papers mainly provided us with an understanding of the difficulties students encounter (beginning level) and a preliminary description of the levels of students' reasoning about social problems. We categorized students' reasoning

into five reasoning activities (e.g., causal analysis) and further divided these into subcategories (e.g., the distinction between long- and short-term causes and consequences). We followed this procedure for each component (describing, explaining, solving), thus constructing three rubrics.

Validation and reconstruction phase

Consulting experts is a standard method for obtaining group consensus on less familiar subjects (Feldon, 2007; Linstone & Turoff, 1975). We conducted two focus groups to validate the rubrics: a teacher panel and a teacher educator panel. The teacher panel consisted of a team of three experienced social science teachers from three different Dutch schools (age 28 to 46 years). The teachers had a master's degree in sociology or political science. These teachers were also test experts working at the Dutch National Institute for Educational Assessment, and have considerable experience constructing social science education tests. We considered this experience relevant because designers have significant expertise in precisely describing which of the students' answers are correct or incorrect, allowing them deep insights into the different levels of student reasoning. The teacher educator panel consisted of four social science teacher educators from four Dutch universities (age 48 to 65 years). They all had a master's degree in sociology or political science. As teacher educators, they had expert knowledge about social scientific reasoning and student learning. Three of the four also worked as social science teachers in secondary education. The two focus groups provided feedback on the list of students' social scientific reasoning flaws and the rubrics through semi-structured interviews.

We divided the interviews into two parts. In the first part, we used an inductive approach by asking the experts to share misconceptions or flaws they often encounter in students' social scientific reasoning. In the second part, we presented the list of students' frequently occurring social scientific reasoning flaws that we found in the papers and asked questions concerning the rubrics' usability, validity, and comprehensibility.

We audio recorded the semi-structured interviews and later transcribed and analyzed them. We used the sociological and social scientific reasoning literature to help us in coding (e.g., Sandahl, 2015; Ultee et al., 2003; van Tubergen, 2020). We began the analysis by "open coding" and identified different patterns emerging from the raw data via constant comparison (Miles & Huberman, 1994). The transcription was first analyzed individually by the first three authors by highlighting the most important issues and then categorizing the usability, clarity, and comprehensibility of the list of flaws and the rubrics. After discussing these findings, we made several revisions to both (see the findings section for the most critical adjustments). The revised rubrics and list of reasoning flaws were validated on the second set of 30 papers. We followed the same cycles as the first set (see exploration phase).

Evaluation phase

Finally, two raters (the first and second authors) checked the reliability of the improved rubrics using inter-rater reliability in an evaluation phase. Therefore, we selected the assignment covering most of the rubrics' subcategories, which was an assignment in which students had to provide policy advice on increasing educational inequality (see Table 1).

Table 1. Features of the assignment on social inequality.

Assignment	Required information	Prior knowledge and skills	Organization
Students investigate whether equal abilities also provide equal opportunities in the Dutch education system. The students investigate this issue using quantitative data and three texts and incorporate the findings into advice to their (self-chosen) political party.	The students use quantitative datasets and three sources selected by the teacher. The teacher provides recommendations for the final assignment, such as 'considering possible undesirable effects of their advice in the short or long term.'	In the preparatory assignments, students map cause-and-effect solution relationships for various sources they study using (given) diagrams, such as "Coleman's boat," to describe underlying mechanisms. In the final assignment, students should also use exam program concepts in their opinions and consider potentially undesirable side effects of proposed solutions.	The preparatory assignments were group assignments (three to four students), the advice (final assignment) was individual, and students were graded. Students had three lessons (45 minutes per lesson) for the assignments.

First, we evaluated the rubrics using five student papers, and differences were discussed. Subsequently, we evaluated a set of 18 papers to determine the inter-rater reliability (scores will be discussed in the results section). These included all papers from this particular authentic learning assignment.

Results

First, we will describe how we conceptualized social scientific reasoning and constructed the rubrics using feedback from the focus groups. Second, we will discuss the results of the evaluation phase in which we used the rubrics to assess the quality of students' reasoning in a specific assignment. Third, we will illustrate the three distinct levels of reasoning with examples from the analyzed student papers. Finally, we will present the reasoning flaws we found in those papers.

Constructed rubrics

We incorporated the five reasoning activities into three rubrics: describing, explaining, and solving social problems. Due to the large size and comprehensive nature of the five reasoning activities, we divided them into subcategories to concretize students' social scientific reasoning. Table 2 gives an overview of the reasoning activities and subcategories. In the construction phase, we decided to integrate the reasoning activity, comparing with the other four reasoning activities; otherwise, in the elaboration of this category, there will be subcategories related to the other reasoning activities. Students are, for example, expected to compare the relative importance of causes (causal analysis), different theories to explain a problem (use of social science concepts, theories, and models), whether sources give the same information (use of source), and compare different perspectives.

The experts in the focus groups recognized the division of the rubrics into three parts (describing, explaining, and solving) as appropriate. However, they displayed differing opinions on the labels of the three levels (in the draft version, we used the labels "naïve," "in development," and "social scientific"). In the focus group with teacher educators, there

Table 2. Overview of the reasoning activities and subcategories per rubric.

Reasoning activity*	Subcategories
Causal analysis	Identifying causes
	Connecting causes, consequences, and possible solutions
	Identifying actors
	Formulating conclusion and hypothesis
Use of social scientific concepts, models, and theories	Reasoning with social scientific concepts
	Reasoning with social scientific theories and models
	Concretizing concepts
Use of evidence	Reasoning with sources
	Reasoning with empirical data
	Critical handling of sources and data
	Comparing and contrasting groups
Use of perspectives and reflection on them	Reasoning with distance
	Using perspectives
	Reflecting on own engagement

*As described above, the reasoning activity "comparing" is a part of each of the four reasoning activities.

was a discussion about the labels of the levels of reasoning. One teacher educator suggested that the label "naïve" was better than "insufficient" or "poor." Others indicated that "naïve" could also have negative connotations. All experts recognized the operationalization of the different levels, particularly that general reasoning at higher levels becomes more nuanced, with increased reservation and a higher tendency toward probabilities than laws. They recommended that the term "nuanced" be operationalized in the rubrics.

For each subcategory described in Table 2, we provided descriptions of students' reasoning divided into three levels: beginner, intermediate, and advanced. In line with the recommendations of the focus groups, we labeled the first level as "beginning;" most students had no training in these reasoning skills, yet they were already relatively confident in their capabilities. We labeled the third level as "advanced." Students in this level would appropriate a disciplinary way of reasoning. We labeled the second level as "intermediate." Table 3 shows an example derived from Rubric 1 ("Describing social problems").

In line with our conceptualization of social scientific reasoning in terms of three components and key activities, specific subcategories were part of all three rubrics. For example, the subcategories "using sources," "reasoning with distance," "comparing perspectives," and "reflecting on own engagement" were part of the rubric that focuses on describing social problems (Rubric 1), the rubric that focuses on explaining the problems (Rubric 2), and the rubric that focused on solving the problems (Rubric 3). While in Rubric 1 "using sources" refers to using sources to support claims about the problem, in Rubric 2 it refers to using sources to support claims about causal relationships. The entirety of Rubric 2 ("Explaining social problems") is attached in Appendix as a detailed example of the structure of each of the rubrics.

Assessment of the quality of students' reasoning in a specific assignment

To check whether we can use the subcategories and levels defined in the rubric to assess the quality of students' reasoning in the context of a specific assignment in a reliable way, two raters (the first and second author) scored a subset of 18 papers. The papers focused on describing, explaining, and generating possible solutions for educational inequality. However, the assignment places most emphasis on reasoning about solutions (students

Table 3. Example of Rubric 1: Describing social problems.

Reasoning activities and subcategories	Beginning	Intermediate	Advanced
The student …		The student …	The student …
Causal analysis: Identifying indicators	Does not identify indicators or confuses indicators of the problem with its causes.	Identifies only one indicator of the problem.	Identifies multiple indicators of the problem *and* identifies these in relation to each other.
… *			
Use of social scientific concepts, models and theories: Using social scientific concepts	Describes the problem mainly as a social problem and uses mostly everyday language.	Describes the problem in part as a social science problem and uses one, or a few, social scientific concepts in doing so.	Describes the problem as a social scientific problem and uses several social science concepts in conjunction.
… *			
Use of evidence: Using empirical data	Does not use empirical data to describe the issue or uses only evidence from immediate surroundings/own experience.	Uses empirical data to describe the issue, but mainly uses evidence from immediate surroundings/own experience.	Uses empirical data to describe the issue and these data are mainly based on (scientific) research.
… *			
Use of perspectives and reflection on them: Reasoning with distance	Explains the problem from their own gut feeling/ emotional point of view.	Explains the problem with some distance (less from gut feeling/emotion).	Explains the issue with sufficient distance.
… *			

*For each reasoning activity, we only included one of the subcategories.

wrote policy advice with possible solutions to counter inequality in education); hence, the subcategories of Rubric 3 (solving) were used as a point of departure. These subcategories were complemented with the unique subcategories of the rubrics focused on describing and explaining. In total, 17 distinctive subcategories were scored (see Table 3).

Consequently, students could receive a maximum of 34 points (0, 1, or 2 points per subcategory). In total, 18 papers were coded. The correlation for the total scores between raters was Pearson's r = .94 with a maximum difference of three points in total. Table 4 shows the Kappa scores per subcategory.

The reliability of the subcategories ranged from k .61 to 1 (with one score of k .61 on the subcategory identifying causes). Four subcategories could not be calculated because all papers received the same score for that subcategory (in practice: 0 points).

Examples of different levels of reasoning activities

We now illustrate the rubrics with the outcomes of our analysis of the 83 papers. We focus on the key reasoning activities.

Causal analysis

Beginning. We labeled causal analysis at the beginning level when students defined a problem superficially and experienced difficulties with reasoning regarding the use of causes and consequences. At this level, students approached problems typically from only one cause and found it difficult to reason with any additional causes. Superficial descriptions of social problems were also illustrated when students confused a problem's indicators with its causes. At the beginning level, students confused actors with causes, were seemingly

Table 4. Inter-rater reliability per subcategory for the assignment on social inequality (n = 18).

Subcategory	Rubric	Kappa
1. Identifying the problem	Describing	*
2. Identifying indicators of the problem	Describing	.82
3. Identifying causes of the problem	Explaining	.61
4. Connecting causes, consequences and possible solutions	Explaining	.65
5. Identifying solutions to the problem	Solving	.79
6. Identifying actors	Solving	*
7. Formulating conclusion and hypothesis	Solving	.89
8. Using social scientific theories and models	Solving	1.00
9. Using social scientific concepts	Solving	.69
10. Concretizing concepts	Solving	.90
11. Using sources	Solving	.89
12. Using empirical data	Solving	1.00
13. Critical handling (of evaluation) of sources and data	Solving	*
14. Comparing and contrasting groups	Solving	*
15. Reasoning with distance	Solving	.66
16. Using perspectives	Describing	.85
17. Reflecting on own engagement	Solving	.67

*Due to no variation, these subcategories could not be calculated.

unaware of any unintended consequences and the relative autonomy of social processes, and interpreted causes and consequences as the conscious actions of individuals. Other examples included students reasoning in terms of laws instead of probabilities and possibilities, regularly confusing causation with correlation, and displaying low levels of caution in their reasoning. In several cases, students confused cause and effect or did not describe how the causes contributed to the problem. See Example 1.

Example 1 "Due to globalization, the Internet is being used increasingly." (Student 321)

In Example 1, the student described globalization as the cause and the use of Internet as the consequence. However, technological development (the Internet) is also causing an increase in globalization. This example illustrates that an oversimplified causal relationship does not explain the connection.

Intermediate. In the intermediate level, students identified more than one cause in their reasoning but only partly described and linked causes, consequences, and possible solutions in a nuanced manner. For example, students described links in a predominantly linear fashion (see Example 2). Moreover, students partially devoted attention to such issues as the direction or strength of the relationship, possible intervening causes, and the distinction between causal links and correlations or self-reinforcing processes. When students formulated a conclusion about possible solutions, but insufficient attention was paid to possible counter arguments, we labeled this reasoning as "in development." See Example 2.

Example 2 "Socialization: A criminal* may be caused by parenting, education, or dealings with other criminals. This makes him more likely to go down the criminal path sooner, which could result in more victims." (Students 4, 11**)

* *Students' original formulation*
** *This paper was a group assignment*

In Example 2, the students explained criminality with some reservation, demonstrated by the words "may cause him." The emphasis, however, was mainly on one direction: criminals are negatively influenced by their environment. However, parenting, education, and interaction with others can generate positive effects.

Advanced. When explaining the problem, the advanced level of causal analysis was reached when students included several causes. A distinction is made between different types of causes, such as different angles (socio-economic, socio-cultural, political-legal), scale levels (micro, meso, macro), and policy-normative vision and/or roles (i.e., incidental and structural causes). At the advanced level, students described links between causes and effects in a nuanced way, such as describing links predominantly in probabilities, recognizing that social developments (trends) are context-specific and changeable, and acknowledging such issues as the direction/strength of connections, possible intervening causes, and the distinction between them. When students discussed one or more solutions, the advanced level of causal analysis was reached when they explained how a solution (possibly) contributes to the preferred situation and how the solution matched the causes of the problem. See Example 3.

Example 3 "Another reason for the number of murders in a society depends not only on income inequality, this also has to do with age and whether a person is living together or not. Around adolescence there is a clear spike in the number of convictions and a person's age. There are several explanations for this, one of them involves a sense of maturity clashing with reality at this stage of life. Young people already consider themselves adults but society still sees them as children. As young people get older, criminal behavior decreases, this is probably because their lives become more stable, they are married have children and have steady jobs. Another possible explanation is that a change in life circumstances increases or does not decrease the likelihood of criminal behavior. Research has shown that when a man is married, he exhibits less criminal behavior than if he is single. However, having children does not play a role in this. Likewise, having steady work also reduces the occurrence of criminal behavior, provided the person is over the age of 26. This connection is also very plausible and strong because a lot of research has been done on it." (students 315, 314, 302*)

* This paper was a group assignment

Example 3 illustrates causal analysis at an advanced level. The links between causes and consequences are described in a nuanced way; the links are described predominantly in probabilities (e.g., "this is probably because . . . ") and consideration is also given to different types of causes (e.g., income inequality, age, living together or not).

Use of social scientific concepts, models, and theories

Beginning. Analyzing students' papers revealed how they encountered difficulties in using social science concepts, models, and theories. At this beginning level, students found it challenging to reason about concepts in a concrete manner. When they described, explained, or reasoned about solutions, students made grand statements, routinely using concepts in an overly simplistic and abstract way. Moreover, we found that students used a limited definition of the concepts. They often used zero, or very few, additional

explanations to clarify their reasoning, which we labeled as "concept dropping." We found little reservation in the reasoning at this first level of social scientific reasoning, nor any explanations of how particular concepts, theories, or models are related. See Example 4.

Example 4 "Criminal behavior can have several causes such as: a poor upbringing, low income, socialization or culture. These circumstances cause them to engage in criminal behavior which leads to problems in society." (students 128, 159*)

* This paper was a group assignment

Example 4 shows that the students named social scientific concepts (e.g., socialization and culture), but they reasoned on a simplistic level about and with these concepts. Moreover, in the follow-up to this quotation, the students did not elaborate any further. Their reasoning with these concepts is limited to "concept dropping."

Intermediate. At this level, we noticed that students could give an accurate description of concepts and partially concretize them. They used appropriate indicators, gave appropriate (counter)examples, or connected the concepts/theories to specific contexts. However, when using these concepts and theories, students tended to use generic (rather than domain-specific) language in their description, explanation, or reasoning about solutions.

Example 5 "Crimes occur in the Netherlands, just as in other countries. However, the way it is punished is not yet working optimally. Seen from the rational actor paradigm, the cause of this problem may lie in the idea that criminals have more opportunity to commit crimes than non—criminals. As was also mentioned briefly in the introduction, the consequence of the sub-optimal punishment of criminals is that people feel unsafe. For example, this is bad for the prosperity of the country, in this case the Netherlands. The solution to this problem may be to punish criminals more severely. By imposing higher sentences on criminals, people are more likely to feel that there is less crime. However, this solution also has a negative consequence. Suspects who are innocent will also receive such high sentences. But, on the other hand, the real criminals will remain under lock and key longer." (students 156, 131*)

* *This paper was a group assignment*

The students in Example 5 used the rational actor paradigm to explain criminal behavior. However, no comparison is made with other social scientific paradigms (such as the conflict paradigm) or theories. We, therefore, classified this reasoning as intermediate.

Advanced. When students used concepts, models, and theories in an advanced social scientific way, they described social problems as social scientific ones and employed several social science concepts. Moreover, students' reasoning was labeled as social scientific when students gave an accurate description of concepts (definition), concretized them by using appropriate indicators and (counter)examples, and specifically contextualized them. Students recognized different theories or paradigms in a given explanation or used different relevant theories or paradigms correctly to explain the problem and discuss the differences. See Example 6.

Example 6 "According to the researchers [of The Spirit Level], there is a relationship between independent variables, such as income inequality and social inequality, and the corresponding dependent variable social problems, including the degree of trust, obesity, and homicide."

6a "This is because the differences in society become greater as inequality grows, leading to more social problems, which you can measure in indicators such as an increase in homicide or obesity [. . .] The level of trust depends on the social inequality within a society."
6b "Source 4 shows that the higher the income inequality is, the more social problems there are, and the less trust there is. The degree of trust thus depends directly on the degree of social inequality. But other intervening variables affect the level of trust." (Students 280, 283, 289*)

**This paper was a group assignment*

Example 6 shows how students used the concept of social inequality on a more advanced level by distinguishing the indicators (6a) and connecting the concept to the specific context (6b). At the same time, the students were cautious in building their reasoning by mentioning that other intervening variables may also be influential (6b). This use of concepts in their social scientific description and explanation was labeled as advanced.

Use of evidence

Beginning. When students failed to provide evidence in their reasoning, or if they only appealed to authority when using it, we categorized their reasoning as beginning. Students' reasoning at the beginning level did not use empirical data to describe or explain the problem. When they used evidence, however, they only tended to do so from their immediate vicinity or based on their own experience; in other words, they demonstrated confirmation bias. Students avoided information and arguments opposed to their views. Moreover, when students did not consider the evidence and data's representativeness, reliability, validity, and generalizability, we labeled their reasoning as beginning. When students compared groups, we found that they primarily focused on the extremes and generalized the data for the whole group. See Example 7.

Example 7 "In the Netherlands, we have recently been dealing with the problem that people in our country no longer feel connected. This is because people are behaving more individualistically. This is to the detriment of social cohesion in our country. To promote social cohesion, we have come up with several solutions." (Students 150, 158 and 120*)

** This paper was a group assignment*

Example 7 illustrates the use of evidence at the beginning level. Students reasoned about the problem of declining social connectedness in Dutch society, but they did not use sources to describe this social problem. Therefore, their statements are not supported by evidence.

Intermediate. Students' use of evidence was labeled as intermediate when students used one or more sources to explain the problem. At the same time, they did not discuss the extent to

which sources support or contradict each other. These students used empirical data to explain the problem, but these data were primarily based on their own experiences. When students compared groups, they were more cautious about making generalizations. Furthermore, students' reasoning at the intermediate level only partly took the representativeness, reliability, validity, and generalizability of the evidence and data into account. See Example 8.

Example 8 8a "In 2014, almost 45% of all welfare recipients were poor, a problem that requires a decent solution as soon as possible (SCP, 2016). This is because welfare recipients who start their businesses or work part-time must give up all the money they earn in return for their benefits."

8b "If the government does not intervene quickly enough, a conflict will arise in no time between the poor and the government. This will not benefit social cohesion in the Netherlands." (Students 3 and 12*)

* *This paper was a group assignment*

Example 8 shows that, in the first part, the students used social scientific evidence (8a) but were insufficiently cautious in reasoning with this evidence. The students focused on only one cause. In the second part, the statements were unsubstantiated and more subjective (8b).

Advanced. We specified students' reasoning as advanced when students used multiple sources to explain a problem and discussed the extent to which sources support or contradict each other. At this level, when using empirical evidence, students used data based on (scientific) research and considered the representativeness, reliability, validity, and generalizability of data. See Example 9.

Example 9 "It has been suggested that the degree of trust in others depends not only on income inequality in a country but also on the degree of individualization in a country. As the research in source 25 shows, the level of volunteering is higher in countries with higher scores on the Hofstede scale. This scale determines the degree of individualism in a country. The higher the score the more individualistic the country. The researchers consider this an indication of the degree of trust in a society in that people only volunteer if they trust the people for whom they are doing it. If no one trusted each other, people wouldn't do anything for anyone else either so there would be less volunteering. You would expect that in an individualistic society there would be less volunteering since people here care less about the environment and more about themselves, the opposite appears to be true. However, if you look at source 24 you will see that the correlation is not that strong as many points are far from the trend line. So there is a correlation it is just not strong and we therefore suspect that there are other causes for this correlation other than just the degree of trust." (students 315, 314 and 302*)

* *This paper was a group assignment*

Example 9 demonstrates the use of evidence at an advanced level: the students used multiple sources (data about volunteering in various countries and a source showing correlations) to explain the problem and discussed the extent to which the sources support and contradict each other.

Use of perspectives and reflection on them

Beginning. At the beginning level, students found it challenging to describe different perspectives on the causes, consequences, and possible solutions to social problems. When students analyzed social problems, they preferred to confirm their perspectives. Students analyzed problems based on their gut feelings and personal emotions. When students were unaware of their biases, rarely reflected on their involvement in the social problem, and routinely confused opinions with facts, we labeled their reasoning as beginning. See Example 10.

Example 10 "People who do receive benefits even though they are not fully disabled should perform social tasks in return for their benefits until they can work normally again. Also, abuse of social benefits must be dealt with harshly. If you commit fraud, you will never again receive benefits, and the amount the fraudster has received must be repaid to the last cent." (student 8)

In Example 10, the student described the use of social benefits with little distance: the problem seems to be described by gut feelings. The student made fierce statements ("if you commit fraud, you will never again receive benefits"), but these statements were not further substantiated. Furthermore, the student used only one perspective to explain the problem.

Intermediate. At the intermediate level, students' reasoning improved when they analyzed the social problem with a certain degree of distance that was based less on gut feelings and emotions. Moreover, students appointed their engagement related to the specific problem. Students described more than one perspective, but social-scientific perspectives and opinions were still used interchangeably. See Example 11.

Example 11 "The school system has been adapted to the average student. All students who are not average are disadvantaged because of this [. . .]. Several solutions are possible to tackle inequalities, and I have chosen the best ones. First of all, it is a good idea if there is customization [. . .]. Second, part of the problem would be solved if more attention was paid to the possibility of progressing from a lower to a higher level. To accomplish this, it would be good to mandate smaller mentor classes." (Student 59)

In Example 11, the student described two possible solutions to contribute to an equal school system. The student described the solutions with distance. However, it is not clear specifically why these two solutions were chosen. The solutions are not supported by evidence. It is possible that the student's personal preferences and personal involvement also played a role. However, this is not clear.

Advanced. At the advanced level, students used multiple perspectives and compared and contrasted those perspectives. They separated opinions from facts, identified their

engagement, and reflected on possible consequences. Students recognized and compared different perspectives on the social problem and explicated those perspectives in their reasoning. In the analyses of the student papers, we did not encounter examples of the advanced level of "Use of perspectives and reflection on them." At the advanced level, students can distinguish their emotions and feelings from other points of view and identify their effect on their point of view. For example, when discussing issues related to globalization, students can remark that the cosmopolitan point of view matches their own. Students should be aware of their own frame of reference and should be able to reflect on this in their reasoning.

Flaws in students' social scientific reasoning

The list of reasoning flaws was essentially created as the first step in constructing the rubrics to arrive at a level description: to define the first level of social scientific reasoning. Nevertheless, we see it as a valuable output of the study because little research has been done on the weaknesses of students' reasoning about social problems.

Based on the analysis of 60 student papers in the exploration, validation, and reconstruction phase, we identified 18 different reasoning flaws (see Table 5). The experts in the focus groups approved the initial list of 17 reasoning flaws. The reasoning flaw "seek confirmation and therefore perceive selectively" was mentioned by the experts in the focus groups, but did not appear in the list we created based on our analysis of the student papers. We then revisited the data and indeed found this reasoning flaw. Subsequently, we added this reasoning flaw to the list of reasoning flaws.

Our analysis showed that, in line with previous research, students make firm statements about social problems but are often unable to substantiate their opinions. There is little reservation in their reasoning, which is illustrated in the quote below:

Table 5. Overview of students' flaws in social scientific reasoning.

	Flaws in students' social scientific reasoning
1.	Students give a superficial description of the problem; they reason from one concept, variable, angle, or cause
2.	Students reason on an overly simplistic level about, and with, social scientific concepts: 'concept dropping'
3.	Students reason primarily in terms of opposites and generalize too quickly
4.	Students confuse the causes of a problem with indicators of a problem
5.	Students confuse simultaneity with causality and correlation with causality
6.	Students overestimate the role actors play in a causal analysis
7.	Students do not consider possible intervening variables when making causal connections
8.	Students formulate causal relationships in terms of laws rather than probabilities
9.	Students do not consider possible unintended consequences of possible solutions
10.	Students cannot explain why a possible solution might work, or explain the underlying mechanism, theory, or model
11.	Students are not attentive to evidence that conflicts with their own conclusion: they look for confirmation of their own views
12.	Students reason with anecdotal evidence rather than with social scientific evidence
13.	Students use the statement as evidence: circular reasoning
14.	Students do not have an eye for different perspectives on the social problem
15.	Students do not recognize their own involvement or frame of reference in describing and explaining social problems, or in suggesting solutions
16.	Students skip research steps; for example, they jump from the research question to (the elaboration of) a research instrument, without considering how a variable can be measured and what indicators for it are
17.	Students reason that if their hypothesis is not confirmed their research must be wrong
18.	Students make false appeals to authority

> Obesity and health conditions are caused by education level. With a high level of education, you get a higher socioeconomic status. With a low socioeconomic status, you have more stress, so more stress eating and low socioeconomic status are associated with unhealthy habits [. . .]. (Students 301, 306, 307)

These students do not reason in terms of the probability of being more stressed if you are less educated, but rather in terms of laws and certainties.

Discussion and conclusion

This study aimed to define students' social scientific reasoning activities in terms of concrete behavior while distinguishing between different levels of quality. Furthermore, it examined whether this conceptualization can be used to assess the quality of students' reasoning in the context of a specific assignment.

In this study, we have empirically validated earlier theoretical conceptualizations (e.g., Sandahl, Ultee, Van Tubergen). In this process, the conceptualization of social scientific reasoning was refined by translating it into five reasoning activities that were operationalized into concrete subactivities, such as "Identifying actors" and "Using social scientific theories and models." In all school subjects, students learn to reason, but the content, concepts, and strategies that students use are often different per school subject or domain. This conceptualization (with concrete reasoning activities) provides a subject-specific elaboration of reasoning, namely within social science education.

Besides this clear conceptualization of social scientific reasoning, we now have language to describe this type of reasoning. This study results in a domain-specific elaboration of civic reasoning and scientific reasoning. This social scientific elaboration is more valuable to teachers; it enables them to apply these insights in their teaching, as opposed to the more domain-generic elaboration of scientific or civic reasoning, which is in line with the recommendation to develop more domain-specific elaborations of civic and scientific reasoning (e.g., Fischer et al., 2014; Goldman et al., 2016; Lee et al., 2021).

In this study, we also described what performance at a particular level looks like so that the quality of reasoning can be assessed. We defined three levels of reasoning for each reasoning activity: beginning, intermediate, and advanced. For example, we operationalized causal analysis at the advanced level when students demonstrated their ability to compare and contrast several causes for identifying a social problem and to distinguish between different types of causes, such as different angles (socio-economic, socio-cultural, political-legal), scale levels (micro, meso, macro), and policy-normative visions or roles (e.g., incidental and structural causes).

Our conceptualization of the beginning level of reasoning was derived from reasoning flaws we identified in students' reasoning in their papers. While there was some research that showed the difficulties adults have in reasoning about social problems—and naïve reasoning modes and flaws in reasoning have been identified in related fields—such insights were lacking for social scientific reasoning. Our study thus contributes to the literature by identifying 18 flaws in students' reasoning about social problems. For example, students tend to reason primarily in terms of opposites or extremes, and they tend to generalize too quickly. These findings are consistent with earlier research on adult reasoning (e.g., Dawson & Venville, 2009; Fiske & Taylor, 2017), historical reasoning (e.g., Stoel et al., 2015), and the

role of causal and epistemic complexity in students' understanding of science (e.g., Perkins & Grotzer, 2000).

Limitations and suggestions for further research

This study is not without limitations. The validation of the rubrics is still limited. First, in this study, the rubrics were validated in a relatively homogenous group of students and on one assignment. Validating the rubric across multiple groups and on multiple assignments would contribute to the definition of a learning progression of social scientific reasoning that can be used to investigate or monitor students' progress. In addition, we rarely encountered examples of social scientific reasoning at an advanced level in the student papers.

Furthermore, we did not validate the rubrics by having teachers use them to, for example, identify flaws in reasoning in student work or to create their task-specific rubrics. Previous research showed that teachers experience difficulties using rubrics for scoring (Brookhart & Chen, 2014; Meier et al., 2006). Consequently, such validation may yield new insights regarding the rubric's usability and clarity for teaching practice.

Follow-up research could focus on a different population, such as first-year sociology students, where there is a greater chance to validate the advanced level of reasoning. In addition, further research could focus on developing educational materials that can be used in a professional development program aimed at teachers' professional growth in teaching social scientific reasoning skills and in developing students' ability to reason about societal problems.

Implications for practice

An important practical implication is that the overview of frequently occurring flaws in students' reasoning and the rubrics can be used to design educational materials and activities to overcome these shortcomings and to achieve a higher level of reasoning. When teachers, for example, become more aware that students sometimes do not consider possible unintended consequences of possible solutions, then they can pay extra attention to this aspect of student learning.

Furthermore, teachers can use the rubrics (with examples of students' reasoning) and the overview of flaws to learn how to diagnose students' reasoning levels. The rubrics and the list of reasoning flaws can give teachers (and students) more insight into students' social scientific reasoning skills and the entry-level of students. These insights can support social science teachers in formulating appropriate learning goals and developing assignments and assessment instruments focused on social scientific reasoning. As the evaluation phase also demonstrated, not all assignments in social science classes will address all three components; sometimes, the emphasis is more on describing and explaining a problem and sometimes more on discussing and weighing possible solutions. Teachers can make selections in the rubrics based on their objectives for the lesson. Therefore, the rubrics can also be a guide for students; it sheds light on what is needed to take reasoning to the next level.

To conclude, this study provides a theoretical model of components of students' social scientific reasoning that is validated by empirical data. These components are structured into a rubric that can be used to evaluate the quality of students' reasoning. This rubric

specifies activities relevant to social scientific reasoning and how these activities can vary across three levels based on empirical evidence from student assignments. The insights from this study can contribute to researchers, teacher educators, and teachers' understanding of social scientific reasoning and what is difficult for students. Developing students' social scientific reasoning is important to support and develop a more robust, deliberative democracy (Abrami et al., 2008; Lee et al., 2021).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Netherlands Initiative for Education Research (NRO) under Grant 4O.5.18540.1O9.

ORCID

Thomas Klijnstra http://orcid.org/0000-0002-2391-6009
Gerhard L. Stoel http://orcid.org/0000-0003-0251-4673
Geerte M. Savenije http://orcid.org/0000-0003-1774-8771
Carla A. M. van Boxtel http://orcid.org/0000-0002-5119-121X

References

Abrami, P. C., Bernard, R. M., Borokhovski, E., Wade, A., Surkes, M. A., Tamim, R., & Zhang, D. (2008). Instructional interventions affecting critical thinking skills and dispositions: A stage 1 meta-analysis. *Review of Educational Research*, *78*(4), 1102–1134. https://doi.org/10.3102/0034654308326084

Acar, O., Turkmen, L., & Roychoudhury, A. (2010). Student difficulties in socio-scientific argumentation and decision-making research findings: Crossing the borders of two research lines. *International Journal of Science Education*, *32*(9), 1191–1206. https://doi.org/10.1080/09500690902991805

Avery, P. G., & Barton, K. C. (2017). Exemplars from the field of social studies education research. In M. M. Manfra & C. M. Bolick (Eds.), *The Wiley handbook of social studies research* (pp. 168–187). Wiley Blackwell.

Bao, L., Cai, T., Koenig, K., Fang, K., Han, J., Wang, J., Liu, Q., Ding, L., Cui, L., Luo, Y., Wang, Y., Li, L., & Wu, N. (2009). Learning and scientific reasoning. *Science*, *323*(5914), 586–587. https://doi.org/10.1126/science.1167740

Barton, K. C. (2011). Wars and rumors of war: Making sense of history education in the United States. In R. Guyver & T. Taylor (Eds.), *History wars and the classroom: Global perspective* (pp. 187–202). Information Age.

Barton, K. C., & Avery, P. G. (2016). Research on social studies education: Diverse students, settings, and methods. In D. H. Gitomer & C. A. Bell (Eds.), *Handbook of research on teaching* (5th ed., pp. 985–1038). American Educational Research Association.

Barton, K. C., & McCully, A. W. (2012). Trying to "see things differently": Northern Ireland students' struggle to understand alternative historical perspectives. *Theory & Research in Social Education*, *40*(4), 371–408. https://doi.org/10.1080/00933104.2012.710928

Brookhart, S. M., & Chen, F. (2014). The quality and effectiveness of descriptive rubrics. *Educational Review*, *67*(3), 343–368. https://doi.org/10.1080/00131911.2014.929565

Cavagnetto, A., & Hand, B. (2011). The importance of embedding argument within science classrooms. In M. S. Khyne (Ed.), *Perspectives on scientific argumentation* (pp. 39–53). Springer. https://doi.org/10.1007/978-94-007-2470-9_3

Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research*, *63*(1), 1–49. https://doi.org/10.2307/1170558

Coffin, C. (2004). Learning to write history: The role of causality. *Written Communication*, *21*(3), 261–289. https://doi.org/10.1177/0741088304265476

College van Toetsen en Examens. (2019). *Syllabus Maatschappijwetenschappen VWO. Centraal Examen 2022*. [Syllabus Social Sciences Education. Central exam].

Dawson, V., & Venville, G. J. (2009). High-school students' informal reasoning and argumentation about biotechnology: An indicator of scientific literacy? *International Journal of Science Education*, *31*(11), 1421–1445. https://doi.org/10.1080/09500690801992870

De La Paz, S., & Felton, M. K. (2010). Reading and writing from multiple source documents in history: Effects of strategy instruction with low to average high school writers. *Contemporary Educational Psychology*, *35*(3), 174–192. https://doi.org/10.1016/j.cedpsych.2010.03.001

Evans, J. S. B. (2002). Logic and human reasoning: An assessment of the deduction paradigm. *Psychological Bulletin*, *128*(6), 978. https://doi.org/10.1037/0033-2909.128.6.978

Evans, R. W., Newmann, F. M., & Saxe, D. W. (1996). Defining issues-centered education. In R. W. Evans & D. W. Saxe (Eds.), *Handbook on teaching social issues* (pp. 2–5). National Council for the Social Studies.

Feldon, D. F. (2007). The implications of research on expertise for curriculum and pedagogy. *Educational Psychology Review*, *19*(2), 91–110. https://doi.org/10.1007/s10648-006-9009-0

Fernbach, P. M., Rogers, T., Fox, C. R., & Sloman, S. A. (2013). Political extremism is supported by an illusion of understanding. *Psychological Science*, *24*(6), 939–946. https://doi.org/10.1177/0956797612464058

Fischer, F., Kollar, I., Ufer, S., Sodian, B., Hussmann, H., Pekrun, R., Neuhaus, B., Dorner, B., Pankofer, S., Fischer, M., Strijbos, J.-W., Heene, M., & Eberle, J. (2014). Scientific reasoning and argumentation: Advancing an interdisciplinary research agenda in education. *Frontline Learning Research*, *2*(3), 28–45. https://doi.org/10.14786/flr.v2i3.96

Fiske, S. T., & Taylor, S. E. (2017). *Social cognition: From brains to culture* (3rd ed.). Sage.

Flemming, R. (1986). Adolescent reasoning in socio-scientific issues: Social cognition. *Journal of Research in Science Teaching*, *23*(8), 677–687. https://doi.org/10.1002/tea.3660230803

Goldberg, T. (2013). "It's in my veins": Identity and disciplinary practice in students' discussions of a historical issue. *Theory & Research in Social Education*, *41*(1), 33–64. https://doi.org/10.1080/00933104.2012.757265

Goldberg, T., & Savenije, G. M. (2018). Teaching controversial historical issues. In S. M. Metzger & L. M. Harris (Eds.), *The Wiley international handbook of history teaching and learning* (pp. 503–526). Wiley.

Goldman, S. R., Britt, M. A., Brown, W., Cribb, G., George, M., Greenleaf, C., Lee, C. D., Shanahan, C., & Project READI. (2016). Disciplinary literacies and learning to read for understanding: A conceptual framework for disciplinary literacy. *Educational Psychologist*, *51*(2), 219–246.

Hogan, K. (2002). Small groups' ecological reasoning while making an environmental management decision. *Journal of Research in Science Teaching*, *39*(4), 341–368. https://doi.org/10.1002/tea.10025

Jacoby, L., & Jaccard, J. (2010). Perceived support among families deciding about organ donation for their loved ones: Donor vs nondonor next of kin. *American Journal of Critical Care*, *19*(5), e52–61. https://doi.org/10.4037/ajcc2010396

Jiménez-Aleixandre, M. P., Bugallo Rodríguez, A., & Duschl, R. A. (2000). "Doing the lesson" or "doing science": Argument in high school genetics. *Science Education*, *84*(6), 757–792. https://doi.org/10.1002/1098-237X(200011)84:6%3C757:AID-SCE5%3E3.0.CO;2-F

Kelly, G. J., Druker, S., & Chen, C. (1998). Students' reasoning about electricity: Combining performance assessments with argumentation analysis. *International Journal of Science Education*, *20*(7), 849–871. https://doi.org/10.1080/0950069980200707

Klijnstra, T., Ruijs, G., Stoel, G., & van Boxtel, C. (2022). *Sociaalwetenschappelijk redeneren* [Social scientifc reasoning]. Landelijk Expertisecentrum Mens- en Maatschappijvakken.
Kuhn, D. (1991). *The skills of argument*. Cambridge University Press.
Kuhn, D. (2007). Jumping to conclusions. *Scientific American Mind*, *18*(1), 44–51. https://doi.org/10.1038/scientificamericanmind0207-44
Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480–498. https://doi.org/10.1037/0033-2909.108.3.480
Lee, C. D., White, G., & Dong, D. (Eds.). (2021). *Educating for civic reasoning and discourse*. National Academy of Education.
Linstone, H. A., & Turoff, M. (Eds.). (1975). *The delphi method: Teaching and applications*. Addison-Wesley.
Lundholm, C., & Davies, P. (2013). Conceptual change in the social sciences. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 288–304). Routledge.
Maddox, L. A., & Saye, J. W. (2017). Using hybrid assessments to develop civic competency in history. *The Social Studies*, *108*(2), 55–71. https://doi.org/10.1080/00377996.2017.1283288
McCarthy Young, K., & Leinhardt, G. (1998). Writing from primary documents: A way of knowing in history. *Written Communication*, *15*(1), 25–68. https://doi.org/10.1177/0741088398015001002
McGrew, S., Breakstone, J., Ortega, T., Smith, M., & Wineburg, S. (2018). Can students evaluate online sources? Learning from assessments of civic online reasoning. *Theory & Research in Social Education*, *46*(2), 165–193. https://doi.org/10.1080/00933104.2017.1416320
McKenney, S., & Reeves, T. C. (2018). *Conducting educational design research*. Routledge.
McNeill, K. L., Lizotte, D., Krajcik, J. S., & Marx, R. W. (2006). Fading scaffolds for argumentation and explanation. *Journal of the Learning Sciences*, *15*(2), 153–191. https://doi.org/10.1207/s15327809jls1502_1
Meier, S. L., Rich, B. S., & Cady, J. (2006). Teachers' use of rubrics to score non-traditional tasks: Factors related to discrepancies in scoring. *Assessment in Education*, *13*(1), 69–95. https://doi.org/10.1080/09695940600563512
Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *The Behavioral and Brain Sciences*, *34*(2), 57–74. https://doi.org/10.1017/S0140525X10000968
Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Sage.
Mills, C. W. (2000). *The sociological imagination*. Oxford University Press.
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
National Council for the Social Studies. (2013). *The college, career, and civic life (C3) framework for social studies state standards: Guidance for enhancing the rigor of K–12 civics, economics, geography, and history*.
Perkins, D. N., & Grotzer, T. A. (2000, April 24–28). *Models and moves: The role of causal and epistemic complexity in student's understanding of science* [Paper presentation]. American Educational Research Association Annual Meeting, New Orleans, LA.
Ruijs, G. J. F., & Klijnstra, T. (2021). *Hogere denkvaardigheden: denkgereedschap voor maatschappijleer* [Higher thinking skills: Thinking tools for social science education. In R. van den Boorn (Ed.), *Handboek vakdidactiek Maatschappijleer* (pp. 169–205). Landelijke Expertisecentrum Mens- en Maatschappijvakken.
Sadler, T. D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. *Journal of Research in Science Teaching*, *41*(5), 513–536. https://doi.org/10.1002/tea.20009
Sadler, T. D., & Zeidler, D. L. (2004). The morality of socioscientific issues: Construal and resolution of genetic engineering dilemmas. *Science Education*, *88*(1), 4–27. https://doi.org/10.1002/sce.10101
Sandahl, J. (2015). Preparing for citizenship: The value of second-order concepts in social science education. *Journal of Social Science Education*, *14*(1), 18–29. https://doi.org/10.4119/jsse-732
Sandahl, J. (2020). Opening up the echo chamber: Perspective taking in social science education. *Acta Didactica Norden*, *14*(4), 1–24. https://doi.org/10.5617/adno.8350
Seixas, P., & Morton, T. (2012). *The big six historical thinking concepts*. Nelson. https://doi.org/10.1016/j.jssr.2015.09.003

Simon, M., & Forgette-Giroux, R. (2001). A rubric for scoring postsecondary academic skills. *Practical Assessment, Research, and Evaluation*, *7*(18), 1–4. https://doi.org/10.7275/bh4d-me80

Stitzlein, S. M. (2021). Defining and implementing civic reasoning and discourse: Philosophical and moral foundations for research and practice. In C. D. Lee, G. White, & D. Dong (Eds.), *Educating for civic reasoning and discourse* (pp. 23–52). National Academy of Education.

Stoel, G. L. (2017). *Teaching towards historical expertise: Developing students' ability to reason causally in history* [Unpublished doctoral dissertation]. University of Amsterdam.

Stoel, G. L., van Drie, J. P., & van Boxtel, C. A. M. (2015). Teaching towards historical expertise: Developing a pedagogy for fostering causal reasoning in history. *Journal of Curriculum Studies*, *47* (1), 49–76. https://doi.org/10.1080/00220272.2014.968212

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, *50*(3), 755–769. https://doi.org/10.1111/j.1540-5907.2006.00214.x

Toulmin, S. E. (1958). *The uses of argument*. Cambridge University Press.

Tytler, R., Duggan, S., & Gott, R. (2001). Dimensions of evidence, the public understanding of science and science education. *International Journal of Science Education*, *23*(8), 815–832. https://doi.org/10.1080/09500690010016058

Ultee, W. C., Arts, W. A., & Flap, H. D. (2003). *Sociologie: Vragen, uitspraken, bevindingen* (3rd ed.). Wolters-Noordhoff.

Van Boxtel, C., Hemker, A., Klijnstra, T., & Ruijs, G. (2017). *Toetsen van denkvaardigheden en conceptuele kennis bij maatschappijwetenschappen* [Assessing thinking skills and conceptual knowledge in social science education]. Landelijk Expertisecentrum Mens- en Maatschappijvakken.

Van Boxtel, C., & van Drie, J. (2018). Historical reasoning: Conceptualizations and educational applications. In S. M. Metzger & L. M. Harris (Eds.), *The Wiley international handbook of history teaching and learning* (pp. 149–176). Wiley.

van Drie, J., & Van Boxtel, C. (2007). 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

van Tubergen, F. (2020). *Introduction to sociology*. Routledge.

Wilkinson, R. G., & Pickett, K. (2010). *Why equality is better for everyone*. Penguin Books.

Wilterdink, N., Van Heerikhuizen, B., Rusinovic, K., & Weenink, D. (2017). *Samenlevingen. Inleiding in de sociologie* [Societies. Introduction to sociology]. Noordhoff Uitgevers.

Wineburg, S. 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. https://doi.org/10.1037/0022-0663.83.1.73

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

Wineburg, S., & McGrew, S. (2016). *Evaluating information: The cornerstone of civic online reasoning*. Stanford Digital Repository. http://purl.stanford.edu/fv751yt5934

Zummo, L., Donovan, B., & Busch, K. C. (2020). Complex influences of mechanistic knowledge, worldview, and quantitative reasoning on climate change discourse: Evidence for ideologically motivated reasoning among youth. *Journal of Research in Science Teaching*, *58*(1), 95–127. https://doi.org/10.1002/tea.21648

Appendix. Rubric 2: Explaining a social problem

Reasoning activity	Subcategory	Beginning	Intermediate	Advanced
		The student …	The student …	The student …
Causal analysis	Identifying causes	Indicates no or only one cause	Identifies multiple causes	Identifies multiple causes and distinguishes between different types, such as: • Perspectives: socio-economic, socio-cultural, political-legal; and/or: • Scale levels: micro-macro; and/or: • Policy-normative visions; and/or: • Roles: such as incidental and structural causes (environmental factors).
Causal analysis	Connecting causes and consequences	Confuses cause and effect or does not describe how the causes contribute to the issue.	Describes links between causes and effects in a partly nuanced way: • Describes links predominantly legal-linear. • Devotes partial attention to issues such as the direction/strength of the connection, possible intervening causes, the distinction between causal links and correlations, or self-reinforcing processes.	Describes links between causes and effects in a nuanced manner: • Describes them predominantly in probabilities. • Identifies that social developments (trends) are context-specific and changeable. • Describes the direction/strength of the connection, possible intervening causes, the distinction between causal links and correlations, or self-reinforcing processes.
Causal analysis	Identifying actors	Does not identify the values, norms, and/or interests of (different) actors in explaining the situation.	Does identify the values, norms, and/or interests of (different) actors in explaining the situation.	Does identify, compare and contrast the values, norms, and/or interests of (different) actors in explaining the situation.
Causal analysis	Formulating conclusion and hypothesis	Does not formulate a conclusion/hypothesis about the main causes	Formulates a conclusion/hypothesis about the main causes, but: • New causes are mentioned; and/or: • There is no consideration of the importance of different causes; and/or: • There is insufficient attention to possible counter-arguments.	Formulates a conclusion/hypothesis about the main causes, and: • This is consistent with previously discussed causes; • There is a consideration of the importance of different causes; • Attention is paid to possible counter arguments, which are subsequently refuted.

(*Continued*)

(Continued).

Reasoning activity	Subcategory	Beginning	Intermediate	Advanced
Use of social scientific concepts, models, and theories	Using social scientific concepts	Uses mainly everyday language in the analysis of causes.	Uses one or a few social science concepts in the analysis of causes. Everyday language still plays a role in the description.	Uses several social-scientific concepts in the analysis of causes.
Use of social scientific concepts, models, and theories	Using social scientific theories and models	• Does not recognize an appropriate theory, model or paradigm in a given statement. Or: • Uses a theory or paradigm incorrectly to explain the problem/issue.	• Recognizes a theory, model or paradigm in a given statement. Or: • Uses an appropriate theory, model or paradigm correctly to explain the problem/issue.	• Recognizes different theories, models or paradigms in a given explanation. And: • Compares multiple theories, models, or paradigms to explain the problem/issue.
Use of social scientific concepts, models, and theories	Concretizing concepts	Gives an incorrect description of concepts (definition) or reasons on a too simplified level with concepts (concept-dropping). That is to say: • Uses no, or incorrect, indicators; and/or: • Does not provide appropriate (counter) examples; and/or: • Does not sufficiently link the concepts/perspectives to the specific context.	Gives an accurate description of concepts (definition) and partly concretizes them. That is to say: • Uses appropriate indicators; or: • Gives appropriate (counter) examples; or: • Connects the concepts/perspectives to the specific context.	Gives an accurate description of concepts (definition) and makes them concrete. That is to say: • Uses appropriate indicators; • Gives appropriate (counter) examples; • Connects the concepts/perspectives to the specific context.
Use of evidence	Using sources	Does not use sources to explain the issue.	Uses one or more sources to explain the issue, but does not discuss the extent to which sources support or contradict each other.	Uses multiple sources to explain the issue, and discusses the extent to which sources support or contradict each other.

(*Continued*)

(Continued).

Reasoning activity	Subcategory	Beginning	Intermediate	Advanced
Use of evidence	Using empirical data	Does not use empirical data to explain the issue or uses only evidence from the immediate vicinity/ own experience.	Uses empirical data to explain the issue; but mainly uses evidence from the immediate vicinity/own experience.	Uses empirical data to explain the issue, and these data are mainly based on (social scientific) research/data.
Use of evidence	Comparing and contrasting data	When selecting data, does not take into account: • Representativeness; • Reliability; • Validity; • Generalizability.	When selecting data, partly take into account: • Representativeness; • Reliability; • Validity; • Generalizability.	When selecting data, take into account: • Representativeness; • Reliability; • Validity; • Generalizability.
Use of evidence	Comparing and contrasting groups	When comparing and contrasting groups, focuses mainly on the extremes and generalizes the data for the whole group.	When comparing and contrasting groups, focuses on the extremes, but is cautious about generalizing the data for the whole group.	When comparing and contrasting groups, has attention for developments across the whole/ middle.
Use of perspectives and reflection on them	Reasoning with distance	Explains the problem mainly from the gut feeling/emotional point of view	Explains the problem with some distance (less from gut feeling/emotion).	Explains the problem/issue with sufficient distance.
Use of perspectives and reflection on them	Using perspectives	• Uses only one perspective to explain the problem, mostly their own • Wants to confirm own perspective • Confuses opinions with facts.	• Uses multiple perspective to explain the problem. However: • Social scientific perspectives and opinions are still used interchangeably.	• Uses multiple perspective to explain the problem • Compares and contrasts the different perspectives • Separates opinions from facts
Use of perspectives and reflection on them	Reflecting on own engagement	Is unaware of one's own bias and does not reflect on own involvement.	Appoints one's own engagement related to the specific problem.	Identifies one's own engagement and reflects on possible consequences for the problem analysis.