International Society of
the Learning Sciences

'Scientists don't know the truth; they tell you what they know to be true!': Shaping High School Students Understanding of How Scientists Establish Trustworthy Claims Using Epistemic Practices

Noora F. Noushad, University of Pennsylvania, noora@gse.upenn.edu
Susan A. Yoon, University of Pennsylvania, yoonsa@upenn.edu
Clark A. Chinn, Rutgers University, clark.chinn@gse.rutgers.edu
Zhitong Yang, University of Pennsylvania, zhitong@upenn.edu
Huma Hussain-Abidi, Rutgers University, hh429@scarletmail.rutgers.edu
Kyle Hunkar, Rutgers University, supposedly.kyle@gmail.com

Abstract: Science skepticism challenges the trustworthiness of scientific knowledge. Researchers suggest that school science curricula should emphasize the epistemic practices real-world scientists use to generate claims, such as actively seeking contradictory evidence for explanatory models and comparing findings with peers. However, empirical evidence supporting the use of epistemic practices, and its potential impact on students' trustworthiness of science remains limited. This study examines four ninth-grade biology students who designed experiments to understand a fictional viral outbreak using agent-based simulation data. They iteratively refined their designs and discussed with peers. Analysis of student worksheets and discussions reveals that students used three epistemic practices: considering multiple explanations, systematically evaluating evidence, and comparing findings with similar experiments. However, they struggled to revise their initial models when presented with conflicting evidence by their peers. These findings offer insights into how students engage with epistemic practices and their perceptions of science's trustworthiness.

Introduction

Science skepticism, characterized by a growing lack of trust in the knowledge claims generated by scientists, has become a concerning phenomenon (Funk et al., 2019; Osborne & Pimentel, 2022). One reason for science skepticism is the mistrust that stems from not accurately understanding how scientists generate trustworthy claims (Bromme & Goldman, 2014; Oreskes, 2019; Sinatra & Hofer, 2021). Common misconceptions about scientific investigations include the belief that scientists always unanimously agree on claims without any conflicting evidence (Sinatra & Hofer, 2021) and that when advancing theoretical assumptions, there aren't practices in place within the scientific community to discourage scientists from engaging in biased practices such as cherry-picking data (Johnson et al., 2023). Consequently, this leads to public doubt when scientists disagree or change their stance on a phenomenon, undermining trust in science (Oreskes, 2019).

One proposal to address these misconceptions is to teach epistemic practices in science inquiry done in schools (Berland et al., 2020; Chinn et al., 2023; Yoon et al., 2023). Epistemic practices refer to the processes scientists use to generate reliable knowledge claims (Chinn et al., 2023; Duschl 2008; 2020). Reliability here refers to claims generated from systematic study of evidence (Oreskes, 2019). These practices involve the social and epistemological decisions that scientists engage in to establish trustworthy claims, such as generating multiple explanations of a phenomenon (Bromme & Goldman, 2014), designing experiments to evaluate evidence that accepts and rejects their initial assumptions of a phenomenon (Chinn et al., 2023), comparing findings with other scientists to confirm the reliability of claims (Duschl 2008; 2020), and revising explanatory models based on emerging evidence and negotiations about how evidence is studied (Chinn et al., 2023). However, this is a challenging task to conduct inside K-12 science classrooms where scientific inquiry is often depicted as a linear sequence of steps that if adhered to often results in one correct answer (Abd-El-Khalick & Lederman, 2023; Davidson et al., 2021; Kite et al., 2021) Often decision-making around what counts as markers of high-quality evidence (Berland et al., 2020; Davidson et al., 2021) and the social processes of comparing findings and resolving disagreements are not emphasized in traditional K-12 science inquiry (Chinn et al., 2023; Duschl 2020; 2021).

Additionally, most inquiry done in schools is too distant from the topics contested in the real world, such as conversations about vaccine safety that occurred during COVID-19. This gap is likely to interfere with students' ability to transfer the reasoning they learn about inquiry to accurately navigate decision-making when science's trustworthiness is contested in public (Chinn et al., 2023). Within the literature on epistemic practices, there are theoretical suggestions on ways to integrate these practices into science instruction through teacher training and curriculum development (Berland et al., 2020; Chinn et al., 2023; Davidson et al., 2021; Manz et al., 2020).

However, there is a lack of empirical data on students' ability to navigate the complexity of using these practices and accurately interpret them as reliable practices to generate trustworthy claims (Chinn et al., 2023). Additionally, we know little about whether their engagement with epistemic practice shapes how they perceive science's trustworthiness.

To bridge this gap, this paper examines how four high school biology students engaged in an inquiry activity where they investigated a fictional virus outbreak, titled the epidemic unit. They did this by designing and conducting experiments using an agent-based simulation modeling tool (Figure 1) and then comparing their findings with those of other students who conducted similar experiments. The goal was to identify which of the four mitigation factors (e.g., masking, distancing, vaccination, or lockdown) could help contain the spread. The curriculum was intentionally designed to enable students to make independent decisions regarding what they considered reliable or unreliable markers of their experimental designs and to reflect on their epistemic practices. In this paper, we address three research questions:

1. How do students use epistemic practices, such as generating multiple explanations, accounting for all evidence, or revising their initial predictions based on new evidence, when formulating their experiments to generate trustworthy claims about the mitigation factors?
2. How did engagement with epistemic practices shape their understanding of science's trustworthiness?

Background

Addressing trustworthiness of science through inquiry

There is contention in the field regarding how science's trustworthiness is addressed in K-12 classrooms. Recently, scholars have highlighted a key issue in science education: the tendency to assert that science is trustworthy without providing a clear explanation of why (Chinn et al., 2023; Kienhues et al., 2020; Tan & Koh, 2023). This criticism is primarily directed at the way scientific inquiry is conducted within school settings.

The Nature of Science (NoS) literature consistently points to a gap between how science inquiry is taught and how it is conducted in the real world (Abd-El-Khalick & Lederman, 2023). Much of this criticism stems from the absence of accurate representation in school science of the epistemic practices involved in scientific investigations. For instance, Cofré et al.'s (2019) systematic review of NoS efforts in the literature revealed that students often find certain elements of scientific investigation, such as the empirical basis, observation, and inference, more accessible to learn than other elements like tentativeness in findings or the social processes that contribute to reconciling findings through multiple comparisons. Furthermore, Kite et al. 's (2021) study involving 125 science teachers found that only a handful of teachers exhibited an understanding of epistemic practices that went beyond the conventional, linear scientific method often depicted in textbooks. They tended to view explanatory models as mere instructional aids for representation and not as evolving representations of what scientists currently believe to be accurate about the phenomenon. They also rarely referenced methodological differences when comparing experiments or the necessity of doing so during inquiry. The literature on NoS emphasizes the need to understand and identify the parts of real-world science inquiry that students find challenging to grasp in order to develop effective instructional strategies that address these issues in classrooms (Abd-El-Khalick & Lederman, 2023; Cofré et al., 2019; Kite et al., 2020).

Epistemic practices of science

We conceptualize science EPs using the **AIR** framework (Barzilai & Chinn, 2017; Chinn et al., 2023). **AIR** suggests that the way people evaluate information is influenced by the **a**im or intent they have when engaging with information (e.g., aiming to reach an accurate conclusion), **i**deals which are the criteria one sets as markers to indicate that they have achieved the aim (e.g., considering multiple opposing explanations and their reasons) and **r**eliable processes which are the actions individuals engage to achieve the ideals (e.g., evaluating multiple sources of information or discerning the quality of evidence cited in articles). We use this framework because the science EPs in the epidemic unit were conceptualized using the AIR framework (Chinn et al., 2023).

When considering science EPs within the **AIR** framework, the EPs can be thought of as **i**deals that can be presented to students as practices scientists engage in the real-world to generate trustworthy claims. These can include practices such as: (1) developing arguments with evidence-based reasons, (2) seek and use high-quality evidence, which can be discerned by variables such as sample size, comparison groups, meta-review etc., (3) expose ideas to critique through peer comparisons and evidence, and (4) evaluate and interpret evidence consistently by modifying claims to fit with discrepant evidence. These were some of the **i**deals shared with the teachers as target epistemic practices to cover in the teaching of the epidemic unit (Chinn et al., 2023; Yoon et al., 2023). When students conduct experiments, they may interpret these **i**deals differently. In other words, they might use various **r**eliable processes to achieve these **i**deals. For example, they could support their claims about the effectiveness of masking by citing specific data on the recovered population. Alternatively, they might suggest that more data is needed to fully understand the effectiveness of masking when other variables such as the infected and dead population are considered.

Inquiry to highlight sciences' epistemic practices

To teach epistemic practices, scholars recommend inquiry environments that represent the messiness of real-world science (Chinn et al., 2021; 2023; Manz et al., 2020; Yoon et al., 2023). For example, Chinn et al. (2021) suggest centering *epistemic unfriendliness*, which involves presenting students with conflicting claims that evoke the same kind of emotional intensity or complexity they encounter outside of school, such as claims about vaccine safety. Additionally, Yoon et al. (2023) advocates using inquiry models that encourage students to understand the challenges of systematically studying phenomena that have inherent randomness, like studying an epidemic in the real world. Lastly, Manz et al. (2020) advocate for inquiry that includes model-test scenarios, allowing students to use real-time data to receive direct feedback on their proposed explanations.

In this study we investigate students' inquiry processes while engaging with a disease epidemic model, we used an agent-based simulation (Yoon et al., 2023) that provided students with real-time feedback on their explanatory models. Students are presented with the challenge of running experiments to determine which mitigation factor is the most effective in containing the viral outbreak modeled in the simulation. They generate multiple hypotheses and use the data they collect on infected, recovered, and deceased citizens (see Figure 1) to generate claims regarding the effectiveness of various mitigation factors. The research team hypothesized that students, as representatives of the broader population, might have developed diverse opinions about various mitigation factors such as masking and vaccination concerning COVID-19, potentially conflicting with one another. Therefore, these variations could be leveraged when generating claims about mitigation factors. Additionally, the curriculum does not prescribe one correct way to generate experimental designs which allows for variation in methodology, thereby emulating the messiness of using epistemic practice in the real-world science inquiry.

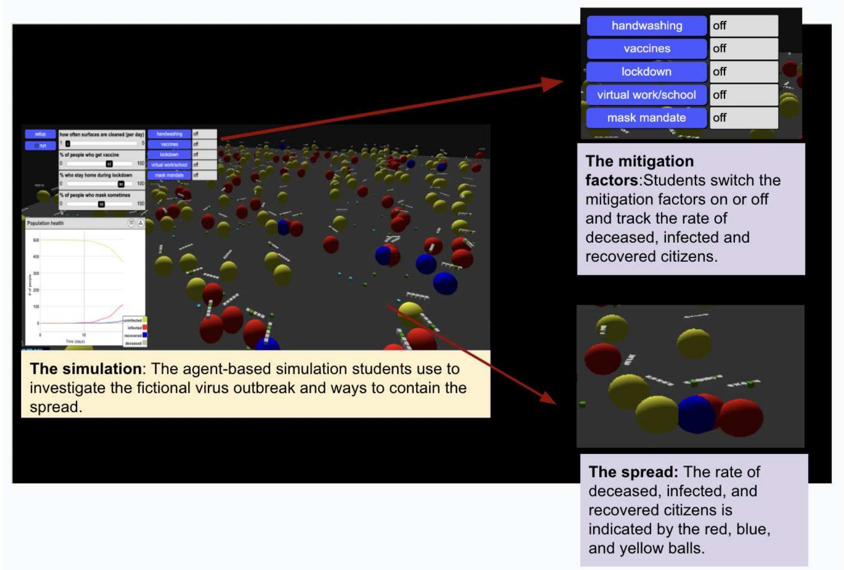
Methods

Curriculum design

This study is part of a larger research project focused on developing professional development for teachers to advance the teaching of epistemic practices in high school science classrooms. The team comprised 12 teachers and approximately 35 students from each of their classrooms spread across India, USA, and Kenya.

As part of the project, the research team developed an epidemic unit consisting of eight 45-minute lessons (Yoon et al, 2023). In the initial two lessons, students negotiated and established a class criterion for reliable epistemic practices they can use to generate trustworthy claims. This criterion involved rules that all student groups must adhere to in terms of designing their experiments. These rules included conducting a minimum of three trials and averaging their results and designing experiments that allowed for comparisons, often by setting up experimental and control groups. During lesson 3, students formulated their research designs. They made decisions about which mitigation factor to investigate and the practices they can use to examine the extent to which their chosen mitigation factor can contain the spread. This included choices regarding the output they wanted to use to compare the effectiveness of their chosen mitigation factor (e.g., deceased, recovered, or infected rates) and the setup of comparison groups to draw inferences about their chosen mitigation strategy, for example comparing deceased rates of masking with deceased rates from distancing. Students were given complete freedom on how they wanted to set up their experiments and asked to justify their reasoning in worksheets. Here students collected preliminary data to inform the predictions they wanted to test. In lesson 4, students executed their experimental designs, they collected data using the simulation and derived conclusions, subsequently comparing their findings with groups that examined similar mitigation factors. In lesson 5, students were introduced to resources to facilitate constructive discussions about differences in findings and ways to reconcile these differences without compromising the findings' reliability. This discussion culminated in a vote to determine the study that generated the most reliable claim during lessons 6 and 7. The vote was determined by evaluating the methods students used to formulate their study and refine their findings with the criterion generated at the beginning of the lesson. The unit concluded with a lesson encouraging students to communicate their findings via a tweet while maintaining transparency about the reliability of the claims. This study analyzed the conversations among four students who investigated the 'masking' mitigation factor during lessons 3 to 5.

Figure 1
The Simulation Students Used to Generate Claims about the Effectiveness of Mitigation Factor in Containing the Spread



Participants

This study evaluated the interactions and reflections of four ninth-grade biology students from a private school in northern India – Amit, Shetal, Shrishti, and Sameer. The student group was chosen based on their decision-making processes during the design and revision process. The decision-making process students engaged in was unique compared to the ones observed in other classes. With this group, the observation was that they were using a strategy different from other groups, who had chosen to set up two conditions to test their claims (e.g., if masking was the chosen factor, then students set up condition 1 with masking on and condition 2 with masking off). However, this group decided to examine evidence in a distinct manner, details of which are reported in the findings section of the study.

All students were in the same biology class and participated in the epidemic unit together. Shrishti served as the group leader, responsible for task management. Shetal navigated the simulation, while Amit and Sameer recorded their reflections on student worksheets. These roles were self-appointed, except for Shrishti, which was assigned by the teacher. The teacher, Ms. Sameera, is a seasoned educator with over 12 years of experience teaching high school biology. This marked her second year of collaboration with the research team. Throughout our planning sessions, she actively contributed by providing suggestions for curriculum design and modifications to ensure the completion of the unit within a one-week duration.

Data sources and analysis

The lead author and another researcher acted as facilitators and data collectors, during the classroom implementation. We collected student worksheets during both lesson 3 and 4 activities and recorded student conversations. These conversations were transcribed to identify episodes where epistemic practices were discussed. This resulted in five videos, totaling about 60 mins each. Each of the five students were interviewed during and after the implementation.

During the inquiry, we posed questions that prompted students to justify the reliability of the practices they used in generating their experimental designs. For example, we asked, "How does comparing results of vaccination with distancing generate trustworthy claims about vaccination?" This approach allowed us to gain insights into the reasoning students applied when making choices about epistemic practices. Furthermore, during post-interviews, students responded to open-ended questions about scientific trustworthiness. These questions included inquiries like, "Has your perception of how scientists conduct investigations changed since the epidemic unit?" and "How do you determine what is true when you come across conflicting scientific information on social media or in the news?" We analyzed a total of 236 minutes of interview data, with an average duration of 77 minutes per interview. These responses helped us discern whether students could connect the epistemic practices they used during the inquiry to the way scientific information is presented in social media. It also provided insights into whether students were reflecting on how the trustworthiness of science is perceived by the general public.

We employed a constant comparative method (Glaser, 2008) to analyze all data sources. This approach helped us understand how students engaged with epistemic practices during their inquiry and the inferences they drew about the reliability of these practices. The analysis proceeded in three phases. First, the lead author inductively identified common themes across the data sources related to how students used and explained epistemic practices (Braun & Clarke, 2012. Second, the authors discussed the coding scheme to define the themes. From this round of coding, multiple themes emerged to classify students' use and understanding of reliable epistemic practices. In this paper, we highlight three most recurring themes: (1) considering multiple explanations, (2) systematic evaluation of evidence and (3) comparing results to similarly run experiments. Examples of coding themes, and data excerpts are presented in Table 1. These themes are discussed in detail in the findings section. Lastly, an external researcher who was not part of the initial coding process audited the analysis. They asked questions and pointed out disagreements (Lincoln & Guba, 1995). We resolved all these issues through discussion until consensus was reached.

Table 1
Examples of the themes with excerpts from Data Analyzed across Student Videos and Worksheets

Themes	Student Data	Rationale
Considering Multiple Explanations	*(Excerpt from classroom video)* Amit: We should go with vaccination, that will be the most effective. Sameer: because of COVID? Amit: Yup Shrishti: What if that's not true for this simulation? Amit: It's most likely to be true for any virus breakout. Shital: But what if this is different, maybe masking will take care of it, and you don't have to get injections. Shrishti: Lets run (the simulation) with just vaccination and then with just masking and see. Students ran the simulation twice and made note of the infected and death rates under vaccination and masking conditions.	Here students initially make predictions about the most effective mitigation factor drawing from their COVID-19 experience. However, one student challenges that these assumptions may not be applicable to the viral breakout under inquiry. So, they proceed to run preliminary trials with vaccination and masking as these were suggested within the group, to make a more-informed prediction backed by the data.
Systematic Evaluation of Evidence	*(Excerpt from classroom video)* Shrishti: hmm (looking at the death rates), comparing masking with vaccination, how can we be a 100% sure that vaccination is better than the others (mitigation factors) we didn't test for? Amit: What if we first run trials without any mitigation factor? That can be our control group. Sameer: Yes, and then we can run each mitigation factor for the same number of days we did with no mitigator factor and then compare each death rate with that of the control group. *(Excerpt from worksheet)* Steps/Rationale "Run 3 trials without any strategies. Record for 40 days/ We can observe the effect of the epidemic on the people and the rate of spread. This would be helpful for further comparison with strategies." "Run multiple trials with each strategy one by one keeping all the other strategies off except the one being tested and record the results. 40 days each/We can see the progress and workings of each strategy to mitigate	Here students run trials with all mitigation strategies turned off and use that as a control group to systematically compare the averaged death rates with trials run with each of the mitigator factors, while the others are turned off. This is an enactment of students actively planning to account for all evidence instead of going with the evidence they collected from comparing death rates from vaccination and masking alone.

	the epidemic and which strategy works the best by seeing which strategy decreases the rate of death.”	
Comparing Results to Similarly Run Experiments	*(Excerpt from classroom video)* Amit: We should go with group 2. They compared vaccination with distancing. Sheetal: Group 3 compared vaccination with masking, we could use them too. Shrishti: hmm. No, but they are both reporting infected and recovery dates. We used death rates. Sameer: We can still compare it with our findings. Let’s use group 2, they ran it for 40 days like us.	Here the group decides to compare their findings with group 2 and 3, they notice there is a discrepancy with the results measured across the selected groups. They decided to go with group 2 as their experiments overlapped more than with group 3.

Findings

Our findings are organized into two main sections. First, we elaborate on the themes, specifically how they emerged during students' inquiries and how students justified the use of these epistemic practices. Next, we present the students' reflections regarding the trustworthiness of science, drawing from their understanding of the epistemic practices involved in generating scientific claims.

What epistemic practices did students use during inquiry: Considering multiple explanations, systematic evaluation of evidence and comparing results to similarly run experiments.

During the inquiry, students demonstrated the use of three epistemic practices: (1) considering multiple explanations (i.e., students actively considered multiple explanations for the effectiveness of different mitigation factors), (2) systematic evaluation of evidence (i.e., students made decisions in setting up experiments that took all evidence into account), and (3) comparing results (i.e., students sought to compare their findings with groups that closely resembled their experimental designs). For example, when discussing the nature of the virus in the simulation, Amit initially suggested vaccination as the most effective strategy, drawing from his experience with COVID-19. However, Shrishti challenged this prediction by questioning its applicability to the current simulation (Table 1). To reconcile their differing views, the group collectively decided to perform a preliminary data examination. They established a control group where all mitigation factors were turned off and set up separate experimental groups for each of the remaining five mitigation factors, isolating one while turning the others off. Their rationale was to systematically compare the average death rates from the control group, where all mitigation factors were turned off, with the rates from trials conducted where every other individual mitigation factor was turned on, sequentially. This decision demonstrated their commitment to consider all available evidence, ensuring a fair assessment of death rates across all mitigation factors rather than designing an experiment that could potentially favor a vaccination or masking. When asked to justify their reasoning, Shrishti commented that “we noticed we were all kind of leaning towards wanting to investigate vaccination because it may be the most effective mitigation factor. If we setup the experiment to compare a condition where vaccination was turned on with a condition where it is turned off, we cannot accurately report vaccinations effectiveness as compared to others” Sameera added, “Yes, it is like saying vaccinations are effective cause we designed our experiment to kind of say it is”. These reflections show that they intentionally constructed an experiment to produce evidence that would challenge their initial hypothesis that favored vaccines over masking. Their reflections also demonstrate a metacognitive awareness of why comparing death rates across multiple conditions (e.g., lockdown, masking, vaccination, and virtual school) was likely to generate a less biased claim.

Furthermore, when comparing findings with peers, they actively sought out groups that had investigated vaccination along with other mitigation factors under similar experimental conditions. Amit commented, "We picked group 2 over 3 because they compared vaccination with distancing and ran the trials for 40 days, so we could compare if their findings matched with ours as the conditions were more similar." When asked how this comparison added to the reliability of their claims, Shital said, "We can see if our results align. If they align, that means our findings about vaccination are likely to be true." These reflections indicate that students were able to understand that comparing results among peers who conducted similar experiments was a practice that enhanced the trustworthiness of their claims. In other words, they were able to engage with and reflect on the value of the social aspect of comparing results across experiments conducted under similar conditions.

However, when comparing their findings with those of their peer group, they noticed a discrepancy. The comparison group found that combining vaccinations and masking resulted in higher recovery rates than using vaccinations as a standalone factor. This presented an opportunity for the students to revise the claim they had generated from their experiment, which initially stated that "vaccinations resulted in lower death rates than other mitigation factors and therefore is the most effective mitigation factor in containing the viral spread." However, the students chose not to revise their claims. When asked to explain their reasoning, Amit said, "They conducted the experiment similarly to us, but they focused on recovery rates while we looked at death rates. That's why our findings differ." Here, Amit accurately acknowledged the methodological difference, but struggled to articulate how to reconcile their initial conclusion about vaccinations in light of conflicting evidence presented by their peers, which suggested that combining vaccinations and masking may be more effective in containing the spread than using vaccinations alone. This indicates that while the students may have grasped the social nature of science, which involves comparing findings as a reliable epistemic practice, they may not have fully understood that scientists compare findings and reconcile differences when they encounter them. Recognizing and reconciling differences is an equally important practice in generating trustworthy claims.

What inferences did students make of science's trustworthiness: "Scientists don't know the truth; they tell you what they know to be true!"

When reflecting on the practices involved in establishing trustworthy claims, students noted the labor-intensive nature of ensuring reliability in science. Amit commented, "This takes a lot of time; we have to run so many trials. Each of us conducted 3 trials to ensure our averages matched." Sheetal added, "We also have to keep comparing with each other." When asked if these practices influenced their perception of scientific trustworthiness, two students offered interesting insights. They acknowledged that the practices of scientists, which rely on evidence and aim to avoid bias, make scientific claims credible. However, they also expressed that scientific claims about new phenomena, such as COVID-19, may not always encompass the complete truth. Sheetal said, "Scientists base their claims on evidence, they analyze data, and they try to ensure their results are unbiased, and they communicate that to others, so yes, it's true, but there can always be something that needs further examination, especially regarding new things like COVID." Srishti added, "We concluded that vaccination is the best based on our data, but the other group found something different. Scientists don't have all the answers, but they share what they believe to be true." In discussing the social aspect of comparing findings, students recognized that scientists have disagreements, but they did not mention the resolution of these disagreements. When probed further about these disagreements, Srishti commented, "it's not like scientists are making guesses, yes, they start with a guess like we did with vaccinations, then they look at the data and tell you what they see, so it is true." These student reflections indicate that they could identify that scientific claims are trustworthy because they are backed by evidence. While they demonstrated an understanding of the epistemic practice of comparing findings and saw value in disagreements, they however, did not exhibit an accurate understanding of the need to resolve them.

Discussion

In the following sections, we discuss the two contributions our findings make to the existing literature on using epistemic practices in inquiry to communicate the trustworthiness of scientific claims. First, we aimed to understand the epistemic practices students used when given the opportunity to design their experiments. Students demonstrated the ability to intentionally set up experimental comparisons to consider multiple explanations and study evidence to evaluate these explanations. This aligns with existing literature findings that indicate students grasp the cognitive dimensions of scientific investigations more quickly (Cofré et al., 2021). Students' reflections on the rationale behind using these practices advance our understanding of the metacognitive awareness students can exhibit when allowed to engage in inquiry with no predetermined correct way to set up experiments (Chinn et al., 2021). However, it's important to note that they encountered difficulties when changing explanatory models while engaging in the social aspect of comparing findings across peer groups. Although students successfully recognized that they had different findings due to the methods they used, their reflections did not demonstrate an awareness that resolving these disagreements is crucial for generating trustworthy claims. Our findings suggest that students may be able to accept productive disagreements among scientists. However, achieving successful reconciliations of these disagreements may be a less accessible aspect, requiring additional instructional support.

Second, our findings on students' reflections about trustworthiness of science provides insights into how students appreciated the evidence-laden nature of scientific investigation while acknowledging the possible limits of scientific knowledge. Students were able to grapple with the often-difficult line between the reliability of science and the limitations of what scientists know. We also noted that students demonstrated a multiplist view of scientific knowledge, which is the belief that varying opinions on scientific matters are equally valid (Sinatra & Hofer, 2021). This may have been due to the emphasis on comparing findings as a reliable epistemic practice and a lesser grasp on the practice of successful reconciliation of the differences as a way to establish a more trustworthy claim. We recognize the limitations of the study with respect to the small sample size examined in this paper and, therefore, do not claim generalization of these findings beyond this particular sample.

References

Abd-El-Khalick, F., & Lederman, N. G. (2023). Research on teaching, learning, and assessment of nature of science. *Handbook of Research on Science Education*, 850-898.

Berland, L. K., Russ, R. S., & West, C. P. (2020). Supporting scientific practices through epistemologically responsive science teaching. Journal of Science Teacher Education, 31(3), 264-290.

Braun, V., & Clarke, V. (2012). *Thematic analysis*. American Psychological Association.

Bromme, R., & Goldman, S. R. (2014). The public's bounded understanding of science. *Educational Psychologist*, *49*(2), 59-69.

Chen, Y. C., & Techawitthayachinda, R. (2021). Developing deep learning in science classrooms: Tactics to manage epistemic uncertainty during whole-class discussion. Journal of Research in Science Teaching, 58(8), 1083-1116.

Chinn, C. A., Barzilai, S., & Duncan, R. G. (2021). Education for a "post-truth" world: New directions for research and practice. Educational Researcher, 50(1), 51-60.

Chinn, C. A., Yoon, S. A., Hussain-Abidi, H., Hunkar, K., Noushad, N., & Richman, T. (2023). Designing learning environments to promote good thinking in a post-truth world: An example from science education. European Journal of Education.

Cofré, H., Núñez, P., Santibáñez, D., Pavez, J. M., Valencia, M., & Vergara, C. (2019). A critical review of students' and teachers' understanding of the nature of science. *Science & Education*, *28*, 205-248.

Davidson, S. G., Jaber, L. Z., & Southerland, S. A. (2021). Encountering uncertainty and learning to "Try and Try and Try": The development of teachers' perseverance within science. In Proceedings of the 15th International Conference of the Learning Sciences-ICLS 2021. pp. 751-758. International Society of the Learning Sciences.

Duschl, R. A. (2020). Practical reasoning and decision making in science: Struggles for truth. *Educational Psychologist*, 55(3), 187-192.

Duschl, R. (2008). Science education in three-part harmony: Balancing conceptual, epistemic, and social learning goals*. Review of Research in Education*, 32(1), 268-291.

Funk, C., Hefferon, M., Kennedy, B., & Johnson, C. (2019). Trust and mistrust in Americans' Views of Scientific Experts. *Pew Research Center*, *2*, 1-96.

Johnson, B. B., Mayorga, M., & Dieckmann, N. F. (2023). How people decide who is correct when groups of scientists disagree. *Risk Analysis*.

Kienhues, D., Jucks, R., & Bromme, R. (2020). Sealing the gateways for post-truthism: Reestablishing the epistemic authority of science. *Educational Psychologist*, 55(3), 144-154.

Kite, V., Park, S., McCance, K., & Seung, E. (2021). Secondary science teachers' understandings of the epistemic nature of science practices. *Journal of Science Teacher Education*, 32(3), 243-264.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. sage.

Manz, E., Lehrer, R., & Schauble, L. (2020). Rethinking the classroom science investigation. *Journal of Research in Science Teaching*, 57(7), 1148-1174.

Oreskes, N. (2019). *Why trust science?* Princeton University Press.

Osborne, J., & Pimentel, D. (2022). Science, misinformation, and the role of education. *Science*, 378(6617), 246-248.

Sinatra, G. M., & Hofer, B. K. (2021). *Science denial: Why it happens and what to do about it*. Oxford University Press.

Tan, M., & Koh, T. S. (2023). Learning to become ignorant: Improving the quality of epistemic knowledge in science education. Science Education, 107(1), 9-27.

Yoon, S. A., Chinn, C., Noushad, N., Richman, T., Hussain-Abidi, H., Hunkar, K., Cottone, A., Katz, J., Mitkus, E., & Wendel, D. 7 (2023). Design principles for teaching complex socioscientific issues: the design of a complex systems agent-based disease epidemic model and the application of epistemic practices in high school biology. *Frontiers in Education, 8*(1),1-14.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 579971. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.