International Society of the Learning Sciences

How Laypeople Resolve Expert Disagreement: The Role of Epistemic Aims and Ideals

Qiuyu Lin, Qianqian Gao, Clark A. Chinn
qiuyu.lin@gse.rutgers.edu, qianqian.gao@gse.rutgers.edu, clark.chinn@gse.rutgers.edu
Rutgers University

Toshio Mochizuki, Waseda University, tmochi@waseda.jp

Abstract: Expert disagreement frequently arises in everyday decision-making, yet little is known about how people's epistemic thinking shapes their disagreement evaluation and resolution. This study applied the AIR framework of epistemic thinking to examine how laypeople's personal and epistemic aims and ideals shape their disagreement resolution strategies. Four triads of undergraduate students assessed conflicting dietary recommendations using a web-based collaborative learning environment, synthesizing the multiple documents into their shared visual graphic organizer to analyze the disagreements. The result showed that students evoked epistemic ideals across three main categories: ideals for authors, for study design, and for documents. The verbal protocol analysis of the students' conversation revealed that the groups followed different paths to deal with the disagreements; depending on their different aims and ideals, each triad came to different conclusions. Our analyses reveal that laypeople's epistemic and personal aims and criteria shape their strategies in resolving disagreement among conflicting documents.

Introduction

People routinely seek guidance from experts when making personal decisions, yet it is common to encounter conflicting expert advice. Expert disagreement is pervasive, stemming both from the inherent complexity of intellectual inquiry and from deliberate attempts to manufacture controversy (Broniatowski et al., 2018). Such disagreements vary widely in their quality and credibility and span a wide range of issues, such as the health impact of artificially sweetened beverages (Hellsten et al., 2010) and the concerns about mammography screening for early breast cancer detection (Bewley, 2011). Exposure to conflicting health advice can confuse individuals and reduce people's trust in science (Chang, 2015). Dietary debates that play out in the media may not only influence short-term dietary decisions but also undermine future efforts to communicate about nutrition broadly (Clark, Nagler, & Niederdeppe, 2019).

Prior research has examined how laypeople understand, evaluate, and ultimately resolve disagreements among multiple conflicting documents (e.g., Barzilai, Thomm et al., 2020; Thomm & Bromme, 2016). While these studies highlight the use of first- (i.e., evaluating the claim validity based on one's own direct knowledge and belief) and second- (i.e., evaluating the claim validity based on evaluating the source trustworthiness) hand evaluation strategies (Barzilai, Thomm et al., 2020; Gottschling et al., 2020), less is known about the underlying epistemic thinking that shapes how individuals evaluate the documents and how these processes guide them moving from evaluation to resolution. In addition, in the context of personal decision-making, individuals also bring in personal aims (i.e., the goals they hope to achieve through decision, such as adopting a diet to lose weight or maintain health) and personal criteria (i.e., the standards they apply when judging the feasibility and sustainability of an action plan, such as whether a diet is easy to follow or clearly explained). Yet little is known about how such personal aims and criteria interact with epistemic thinking and further shape resolution.

This study addresses this gap by investigating disagreement resolution within a computer-supported collaborative learning environment in which students discussed disagreements among the documents related to a health topic and unpacking this process through the lens of the AIR framework of epistemic thinking. Specifically, we examine how learners' personal and epistemic aims and ideals emerge when evaluating conflicting documents, and how these different ideals work together to support their decision-making processes.

Theoretical framing and review of the literature

Engaging with expert disagreement in multiple conflicting documents

When people engage productively with expert disagreement in multiple conflicting documents, they may engage in three interrelated processes (cf. Barzilai & Thomm, 2023): (1) evaluating information validity (i.e., first-hand

 © ISLS

evaluation) and source trustworthiness (i.e., second-hand evaluation) to decide "what is true" and "whom to believe" (Barzilai, Thomm et al., 2020; Stadtler & Bromme, 2014); (2) employing research-process- or expert/source-related reasons to explain, regulate, and resolve conflicting claims (Stadtler & Bromme, 2014; Thomm et al., 2015); (3) integrating information across multiple trustworthy documents to construct reconciled intertext and integrated situational model (cf. Barzilai & Thomm, 2023; Rouet & Britt 2011). Prior studies have shown that students employ a wide range of resolution strategies, ranging from ignoring the conflict and merely accumulating factual details to weighing and weaving ideas across multiple documents (Stadler & Bromme, 2014). Evaluation is central to this process (Gottschling et al., 2020). Importantly, people often engage in evaluating multiple aspects of the documents, such as the source trustworthiness and expertise (Gottschling et al., 2020); the validity of the information itself, discourse features in the document, and its corroboration with other available information (Barzilai, Thomm et al., 2020). Collectively, this body of work has clarified the role of evaluation in disagreement resolution and highlighted that the evaluation is grounded in diverse criteria rather than a single universal standard.

However, three important gaps remain. First, although readers' integration criteria have been shown to influence disagreement resolution in multiple conflicting documents, there is still limited understanding of how readers establish and apply these integration criteria when resolving disagreements (Barzilai & Strømsø, 2018). By integration criteria, we refer to the standards used to evaluate the final products of integration. In our study, this specifically denotes the epistemic ideals applied to the conclusions. When integrating multiple documents, readers construct a task model (i.e., a cognitive representation of goals, standards, and strategies for accomplishing the task) (Rouet et al., 2017), which in turn shapes their different strategic approaches to develop conclusions (Bråten et al., 2011). This highlights the central role of readers' aims and the criteria they adopt in integrating multiple (conflicting) documents. In practice, people appeal to a range of reasons for constructing knowledge from multiple conflicting sources, such as accumulating a large amount of information, identifying the source with trustworthiness, and collecting multiple perspectives (Barzilai & Zohar, 2012; Thomm et al., 2017). We argue that these reasons function as ideals for what counts as adequate conclusions. Yet the role of ideals for conclusions in shaping disagreement resolution strategies remains underexplored.

Second, existing studies often underestimate the complexity of epistemic thinking in evaluating and resolving disagreements. There is no single, universal criterion for defining a "high-quality" document and conclusion. Instead, readers rely on multiple, and sometimes competing standards, such as authority, expertise, scientific evidence, perspectives, or bias (Barzilai & Zohar, 2012). Lay readers must therefore coordinate diverse criteria in order to evaluate and reconcile multiple conflicting accounts. However, systematic evidence is lacking on how students configure and coordinate first- and second-hand evaluation (Barzilai, Thomm et al., 2020), or how they balance criteria for sources, content, and methodology. Additionally, some studies highlighted the degree of expert consensus as a hallmark criterion for resolving disagreements (Bromme & Goldman, 2014). Evaluating consensus can involve multiple strategies, such as identifying agreements articulated by scientific associations or assessing review studies that synthesize findings on a given topic. Yet, little is known about how students actually apply consensus-based criteria in practice.

Third, little is known about the mechanism through which first- and second-hand evaluations shape subsequent integration and resolution strategies. Evaluating documents does not always lead to a single "winner" that can be directly endorsed (Thomm & Bromme, 2016). In personal decision-making contexts, individuals may consider personal aims and criteria for action plans when determining how to resolve disagreements. In case of genuine expert disagreement (as opposed to manufactured disagreement), readers often confront multiple reliable but conflicting documents, raising the question of how ideals for documents and for s guide the uptake, ignorance, combination, and interrelation of texts in multiple documents.

Recent efforts have begun to address these issues. For example, Barzilai and colleagues (2020) developed digital epistemic scaffolds that support evaluation and integration by facilitating collaborative document mapping, which enables readers to articulate and elaborate criteria and processes for both documents and the final conclusions. Building on this line of work, the present study seeks to extend our understanding of how people develop and coordinate multiple ideals for sources, content, and conclusions, and how these ideals work together to shape fine-grained strategies for resolving expert disagreement.

Applying the AIR framework to understand epistemic thinking in resolving disagreement

While epistemic thinking plays a crucial role in evaluating and integrating information (Barzilai & Zohar, 2012; Bråten et al., 2011), prior research has predominately approached it from two perspectives: the developmental model (e.g., Barzilai & Zohar, 2012; Thomm et al., 2017) and the epistemic beliefs perspective (e.g., Bråten et al., 2011). Epistemic thinking varies considerably across individuals and contexts when evaluating and integrating

sources (reviewed by Barzilai & Zohar, 2012), which could not be fully captured by perspectives that conceptualize epistemic thinking as a general and stable disposition. Therefore, our study focuses on documenting epistemologies in action during disagreement resolution in multiple conflicting documents. We adopt the AIR framework of epistemic thinking (Chinn & Rinehart, 2016) to analyze students' goals, criteria, and strategies adopted. Compared with other perspectives, the AIR framework provides a more flexible account of this variability across different stages of the task. It enables us to trace how aims and ideals emerge in situ, while also providing a bottom-up and fine-grained lens for characterizing the complexity of students' epistemic reasoning (Barzilai & Chinn, 2024). This framework has three interdependent components: (1) epistemic Aims: these are goals people set, related to inquiry and finding things out. For example, people want to make a science-informed decision on the best diet. (2) epistemic Ideals: these are the standards or criteria used to judge whether an aim has been achieved, and to assess the quality of the related knowledge products. For example, laypeople may assess whether a source is trustworthy using criteria such as expertise, integrity, and lack of bias. (3) Reliable processes: these are the strategies used individually or collaboratively to achieve epistemic aims in light of the relevant ideals. For example, laypeople might verify a source's credentials and consider potential biases as a strategy to identify information from trustworthy sources.

In this study, we examine how the three components of epistemic reasoning interact during group engagement with conflicting documents. In particular, we explore how students articulate aims and ideals, and how these, in turn, shape their disagreement resolution strategies. We treat students' disagreement resolution strategies as forms of epistemic processes. Additionally, in the specific task context of drawing conclusions about diet, we also attend to students' personal goals and standards relevant to actionable decisions about diet. Our research questions are the following: (1) What epistemic ideals are used when undergraduates engage in the multiple documents task with conflicting documents? (2) How are students' aims and ideals related to their disagreement resolution processes? (3) To what extent do students use expert consensus and reviews of empirical studies to resolve disagreement? We approach this problem through fine-grained qualitative analysis to capture how three components of epistemic thinking come together in students' group discussions.

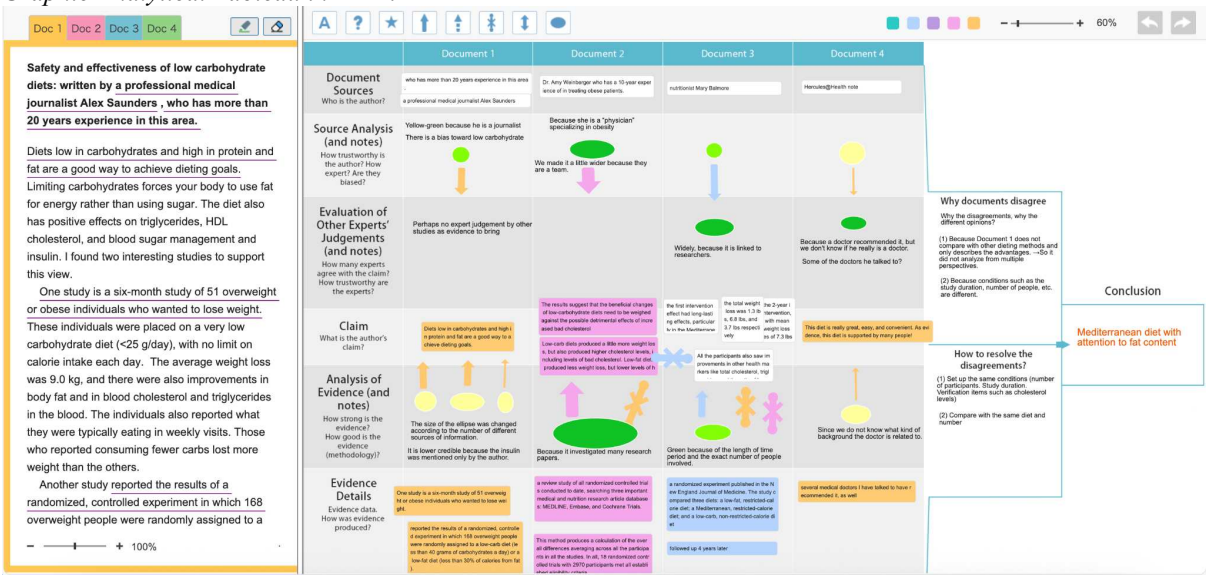
Methods

The CSCL platform and task design

Mochizuki et al. (2022) developed a computer-supported collaborative learning system called EDDiE. In this study, students in each triad used the system to evaluate conflicting documents through a graphic analytical tableau (see Figure 1 for an example) grounded in the Grasp of Evidence framework (Duncan et al., 2018). The analysis tableau's framework serves as an epistemic scaffold (Weinberger et al., 2005) that guides the thinking processes of (1) reading & evaluating documents; (2) reasoning disagreement resolutions; and (3) reaching reasonable conclusions as resolutions for the disagreements.

Students first comprehended and analyzed four documents, each of which explains the effectiveness and/or risks of three types of diet (i.e., low-carb, low-fat, and Mediterranean diet) with different levels of evidence, ranging from solid review studies to no evidence at all. Three of the four documents come from experts in different areas and have other nuance differences, while one document was an easy blog post by an anonymous author. After reading these documents, students first collaboratively extracted, analyzed, and recorded important elements (e.g., sources, claims, evidence) of each document in the analysis tableau. Next, they evaluated each element using graphic symbols. Taking evidence evaluation as an example, they created arrows with differing thicknesses to represent how strongly the corresponding evidence supports the claim, and with differing colors to represent which position the evidence supports. Moreover, students evaluated the quality of evidence while discussing the appropriate color or size of evidence circles: high-quality evidence is colored dense green, poor-quality evidence is colored yellow, and the size of the circle reflects the quantity of evidence. As such, students evaluated the quality and strength of each of these elements in each document. They then identified consistencies or inconsistencies of multiple elements across the documents with the help of various graphic symbols. Afterward, students in each group identified possible reasons for the disagreement. Finally, based on their analysis and group discussion, students were asked to suggest ways to resolve the disagreements and then wrote shared conclusions about the best diet.

Figure 1
Graphic Analytical Tableau in EDDiE



Data sources and analyses

This paper draws on data from a Zoom-based study involving twenty triads of undergraduate students. Each group's conversation was video-recorded and transcribed. The collaborative discussion lasted approximately 2 hours. Four groups were randomly selected from the full dataset to support exploratory analysis and coding scheme development. Transcripts were segmented into analytic units based on shifts in students' interactional engagement with a specific epistemic task (e.g., identify and evaluate sources in each document). Within segments, coding was conducted at the level of meaning units, allowing multiple codes per segment. Transcripts were analyzed sequentially to preserve the trajectory of the epistemic thinking in each group. We first developed initial codebooks from bottom up and top down. Then, two researchers independently coded one group, reviewed and discussed all codes, and refined the codebook based on these discussions to develop a more nuanced understanding of the study context. Following this, two researchers independently coded the remaining three groups, discussed any lingering coding questions, and refined the codebook to develop a shared understanding of the coding procedure and rationale. To make sure the updated codes were accurately applied to each group, two researchers independently re-coded all four groups based on the updated codebook, then met to discuss the disagreements and the emerging patterns. We developed an iterative coding scheme consisting of 38 codes capturing epistemic ideals for trustworthy authors (n = 7), good documents (n = 11), good study designs (n = 6), and good conclusions (n = 8); as well as personal ideals for diet (n = 6). Given the variability in how different groups resolve disagreement, the epistemic processes for disagreement resolution were analyzed as emergent patterns rather than a fully specified coding schemes. Representative codes are illustrated in the Finding section. We established analytic rigor through repeated joint coding, iterative discussion to resolve disagreements, and ongoing refinement of coding schemes. This approach emphasizes interpretive depth and analytic development. Based on the coding results, the research group further conducted verbal protocol analysis (cf. Chi, 1997) for all four groups. Together, we represented each group's epistemic thinking by examining each code and category, mapping the relationships among codes, and developing the symbol representation of each group's epistemic thinking. Two theoretically informative cases are presented to illustrate some common epistemic configuration paths observed across groups.

Findings

RQ1: What epistemic ideals are used when students engage in this multiple-documents task with conflicting documents?

Our analysis shows that students drew on a wide range of epistemic ideals when engaging in both first-hand and second-hand evaluation. Across the four groups, students' ideals fell into three main categories: (1) ideals for trustworthy authors, (2) ideals for good documents, and (3) ideals for good study designs. Rather than

systematically comparing groups, the key finding is the variety of epistemic configurations that emerged within group discussions.

When evaluating the credibility of authors, students often adopted applied several ideals together when assessing the credibility. For example, in Group 2, student C remarked, "He has 10 years of experience treating obese patients, so I think he's quite reliable," drawing on both length of experience and expertise. Student A built on this by adding, "Ten years as a doctor is better than 20 years as a journalist," foregrounding clinical experience as a more salient ideal than professional experience in a relatively unrelated field. Later, student C extended the reasoning by emphasizing collective expertise: "A research team means multiple people who have been working on obesity treatment for a while. That might be very trustworthy." This episode illustrates how students not only applied different ideals for authors but also negotiated their relative importance (e.g., clinical experience is more important than length of experience), a pattern observed across groups.

Guided by the epistemic scaffolds, students also evaluated the quality of documents based on their empirical support and the amount of evidence provided. Additional considerations included a persuasive balance of perspective, accuracy, and specificity of reporting, generalizability of the claim, and so on. For example, one student remarked, "It's written in this wording, but there are no numbers, so I wonder if the numbers will be quite important... It doesn't say much about how much effect it has on cholesterol and cardiovascular risk," emphasizing the importance of concrete numerical data or comparable quantifiable results, so that they could make sense of the conflicting results across documents. These practices indicate that the undergraduates do not rely on a single criterion (e.g., empirical support) but instead take different factors into consideration and mobilize a repertoire of ideals for what counts as a "good" document.

Another recurring category concerned the study design and the methodological soundness of studies. Students adopted criteria such as sample size, sample diversity, systematic data collection method (e.g., randomized control trials, review study with explicit inclusion criteria), the duration of observation, and so on. As we can see in the table, students adopted relatively more ideals in evaluating authors and documents, while less in study design. However, this category of ideals, especially sample size and the duration of observation, becomes decisive when students decide the trustworthiness of each source.

Our scaffolds provided no prompts targeted at what constitutes "good conclusions." As a result, ideals for conclusions were rarely made explicit. However, students' practices, such as only integrating information from trustworthy sources, implicitly revealed their expectations for what makes strong conclusions. We further elaborate on this point when addressing RQ2.

RQ2: How are students' aims and ideals related to their disagreement resolution processes?

We draw on Groups 2 and 3 as examples to examine how different personal and epistemic aims and ideals were configured in practice, and how these different configurations shaped distinct disagreement resolution processes. Figures 2 and 3 depict the group's epistemic thinking processes, highlighting the interplay of aims, ideals, and reliable processes as they worked to resolve disagreements in the multiple conflicting documents task.

Group 2

As shown in Figure 2, Group 2 coordinated a mix of epistemic aims and ideals when resolving disagreement. Specifically, they drew simultaneously on ideals for sources and ideals for conclusions, using this combination to develop different strategies for managing conflicting claims. By "ideals for sources," we refer to the criteria students applied in evaluating the four conflicting documents, including ideals for authors, study designs, and the documents themselves. In the following, we trace Group 2's epistemic thinking process in detail.

Compared to other groups, Group 2 not only identified differences across documents as possible reasons for disagreement but also engaged in systematic evaluations of each source's credibility. Their temporally overarching ideal for strong conclusions was that it should be grounded in credible sources. To enact this ideal, they engaged in a set of five ideals for good sources, which also guided two reliable processes for resolving disagreements: evaluating the credibility of each source, and dismissing lower-credibility documents while retaining those judged as stronger.

First, students emphasized accurate scientific reports. For example, they compared the result reporting in Documents 2 and 3: "the ratio (in Document 3) is easier to understand than the numbers (in Document 2)," and "maybe the actual ratio is what matters." Here, they argued that the ratio between good and bad cholesterol was more meaningful than absolute increases or decreases in cholesterol levels. Thus, they privileged documents that reported results in a way that could be interpreted meaningfully. Second, they highlighted empirical support, noting that "Document 1 does mention the research to a certain extent. Document 4 does not mention it, so it's not good." In their view, documents explicitly referencing supporting studies carried greater weight. Third, the

students drew on expertise as a criterion. “The second one is reliable because it’s a doctor who is speaking. But in the fourth one, we don’t even know if the person is really talking to a doctor.” Here, clinical expertise was treated as a critical marker of credibility, which led them to value Document 2 over Document 4. Fourth, they attended to study design, reasoning that high-quality studies should include larger sample sizes and longer periods of observation. These considerations favored Documents 2 and 3.

In total, the group articulated five ideals for evaluating sources. According to their analysis, Document 2 met three out of five criteria, and Document 3 met four criteria. Documents 1 and 4 met fewer and were subsequently dismissed: “I haven’t really looked at (Documents) 1 and 4 anymore.” In this way, the group operationalized their ideal for strong conclusions by systematically narrowing their focus to Documents 2 and 3, which they judged to be trustworthy. Crucially, no single ideal was sufficient for dismissal on its own. For instance, although the criterion of accurate reporting led them to favor Document 3 over Document 2, it did not prompt them to reject Document 2 entirely. Instead, the group combined multiple ideals, considering how each document measured up across different criteria of trustworthiness. The interplay between ideals and reliable processes enabled them to support their conclusions on trustworthy sources.

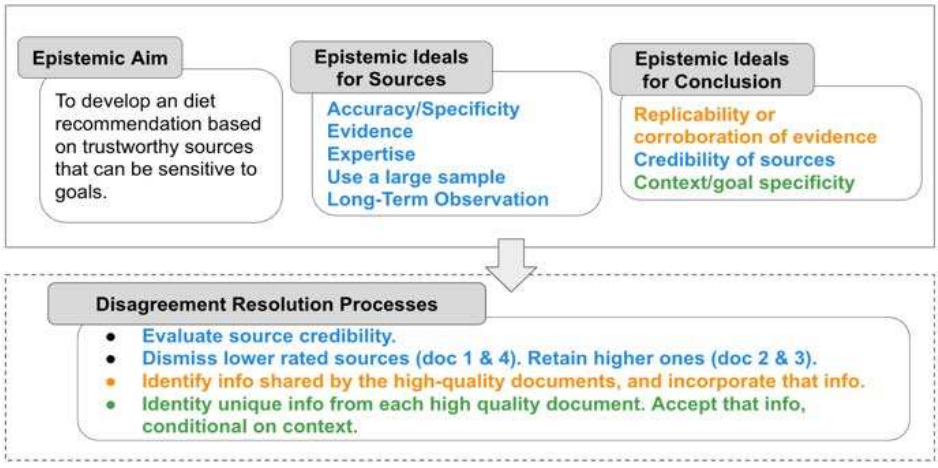
This reasoning led Group 2 to articulate a second ideal for strong conclusions: corroboration of evidence. To pursue this ideal, they adopted the reliable process of identifying information shared by high-quality documents and incorporating it into their conclusions. For example, they aimed to identify agreement across trustworthy documents to make their conclusions more convincing. They reasoned: “There are a few differences in perspective between Documents 2 and 3, but either way, Documents 2 and 3 are basically saying that low-carb diets aren’t that good, right?” Because both documents pointed in the same direction, the students reconciled the discrepancy through the explanation of “differences in perspective” and treated the agreement as a basis for strengthening the credibility of their conclusions.

Building on this, they introduced a third ideal for strong conclusions: context and goal specificity. This ideal was enacted through the reliable process of identifying unique information from each high-quality document and accepting it depending on different contexts. As one student explained, “In addition to this conclusion, I think it will vary depending on each person's constitution, so should we choose the one that is better for us?” Here, the group acknowledged that unique information from different trustworthy documents could be valid but only when applied to the right health goals and individual conditions.

Together, these ideals guided the group to nuanced conclusions. They combined agreement across trustworthy documents with unique contextualized information, resulting in conclusions that is both evidence-based and flexible for different people. “If your only focus is on weight and you want to slim down in a short period, go for a low-carb diet. If you want to lose weight in a healthy, long-term way, choose the Mediterranean diet. Choose a method that suits your body type. Consult with a trusted doctor.” Their conclusions reflected the coordinated use of multiple ideals and reliable processes.

Figure 2

Group 2: The Configuration of Epistemic Aims and Ideals and its Connection with Disagreement Resolution Processes



Group 3

Group 3's epistemic aim and criteria led them to different disagreement resolution strategies. Unlike Group 2, they did not settle on a process for weighing the credibility of sources. Instead, they initially focused on two ideals for sources: (1) consistency of research conditions; (2) persuasive balance of perspective.

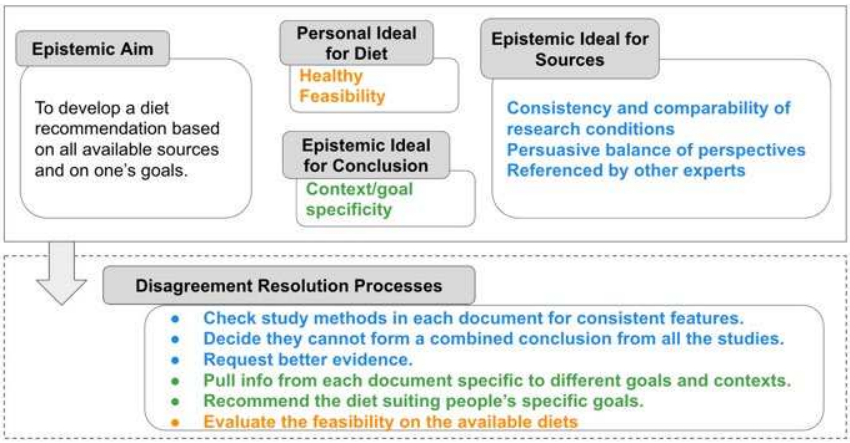
Students reasoned that strong documents should present studies conducted under comparable conditions, such as similar study duration and consistent variables, so that findings can be fairly interpreted without confounding differences. As one student noted, "we might have to look at studies with consistent conditions again—or actually see them ourselves." In addition, they emphasized that a persuasive document should present both supporting and opposing perspectives, while also explaining why one side is more convincing. For example, students in Group 3 thought out loud about the process of developing better documents, "We have to bring them without looking at the results," and "I'm thinking that instead of just looking for evidence to support your own argument, you also need to find evidence that opposes it. By doing that, you can show that there are opposing opinions, but your opinion is stronger because it's more convincing." When students applied these two criteria to the four documents and checked the study methods in each document for consistent features, they found that all four documents failed to meet the criteria. The studies varied in conditions, and the documents did not present balanced perspectives. So the group concluded that they could not form combined conclusions from all the studies. Their proposed solution was to generate or gather new data under consistent conditions and to present the results with balanced perspectives.

At this point, to address the task of coming up with conclusions, the group shifted to an ideal for conclusions: goal and context specificity. Rather than applying any ideals or strategies to distinguish between more and less trustworthy sources, they treated all documents as equally valid. They extracted information from each document and organized it according to different health goals and contexts. One student put it this way: "When asked what the best diet is, I feel like it's not suitable for everyone." They decided to recommend the diet according to specific goals and health conditions mentioned across various documents.

Their reasoning was also shaped by their personal ideals, indicating that diets should be feasible and healthy. They intentionally attend to the feasibility of diets, although there's no text indicating this aspect. They commented, "The Mediterranean diet seems to be the easiest to stick to." They therefore leaned slightly toward the Mediterranean diet because it aligned best with their own sense of what a "good diet" should be. Ultimately, their conclusions reflected their configuration of different ideals: "It is necessary to choose a method according to your goal. A low-carb diet is suitable for short-term weight loss. However, it is not safe for people with diabetes. For long-term goals, the Mediterranean diet is more appropriate. It has fewer food restrictions, making it easier to continue. Additionally, it provides well-balanced nutrition." In short, Group 3's disagreement resolution process relied less on discriminating between more and less credible documents and more on tailoring information from multiple sources to specific contexts and personal ideals.

Figure 3

Group 3: The Configuration of Epistemic Aims and Ideals and its Connection with Disagreement Resolution Processes



RQ3: To what extent do students use expert consensus and reviews of empirical studies to resolve disagreement?

Although the degree of expert consensus is a hallmark criterion for final conclusions, and examining the results of a review study represents a reliable process for resolving disagreements, participants adopted different ideals

to further evaluate the trustworthiness of the review study. Their evaluation reflected meaningful sourcing practices but also revealed a limited awareness of the epistemic value of the review study in decision-making.

Three groups judged the review study as somewhat trustworthy primarily because of its large sample size. For instance, Group 1 emphasized that the study included 18 randomized controlled trials but treated this as equivalent to simply having a large sample, without recognizing the difference between a single study with many participants and a systematic review. Other groups adopted ideals for study design as markers of credibility: strict inclusion criteria, systematic paper selection, clear reporting of data collection and results, and the use of major medical and nutritional databases. For example, Group 3 commented, “Since they disclose the databases they've researched, maybe more people are looking at it.” “The study period is well-defined, and the search criteria are narrowed down properly.”

Groups also raised concerns about the review. Several groups paid attention to cherry picking. Group 2 and 3 dismissed the review as “but it's just a little bit of stuff in there based on different subjects” or “just gathered a bunch of research papers and put them all together.” Group 3 further questioned database selection, noting that “there are so many possibilities, so it ultimately comes down to which site you're looking at.” Group 4 expressed doubts about generalizability, worrying that overly specific and strict inclusion criteria could limit generalizability, “the author, as a doctor in the field, so I'm sure there are people in all kinds of... conditions. But in such a detailed and particular way... I'm afraid that if the conditions were different, I'd say no.”

Overall, participants demonstrated thoughtful sourcing practices by adopting various ideals to evaluate features of the review study and raising valid concerns. However, none of the groups highlighted the key features of the review study: its result reflects large bodies of evidence. Similarly, although some groups integrated the argument presented by the Japanese Medical Association for Diabetes in their final conclusions, none of the groups emphasized the consensus presented by the association as a marker of broader expert agreement, and therefore a convincing reason for relying on their claim.

Discussion and conclusion

Our study showed that students integrated information from diverse documents, but their approaches and conclusions varied depending on the aims and ideals they prioritized. This underscores that adopting different configurations of ideals in relation to the context is central for disagreement resolution processes.

Through the lens of the AIR framework (Chinn & Rinehart, 2016), this study describes how students dynamically adopted both personal and epistemic aims and ideals to the task context. Consistent with prior work on variability in epistemic thinking (Barzilai & Zohar, 2012), our findings provided detailed evidence to show how students flexibly drew on first- and second-hand evaluation strategies (Barzilai, Thomm et al., 2020) through the adoption of a repertoire of ideals in different categories. While existing research highlights the role of source-based criteria such as expertise and vested interests (Bromme et al., 2015; Goldman, 2001), our studies reveal that students also developed ideals for document content and study design when sources alone were insufficient. Moreover, we extend this repertoire by identifying ideals for conclusions, that is, the integration criteria, which have received limited attention in previous disagreements resolution research but proved crucial for resolving conflicts in our task.

This study also extends current understandings of the complexity of epistemic thinking in resolving disagreement in multiple conflicting documents, by depicting both the configuration of epistemic ideals and the process from evaluation to resolution. Prior research indicates that judging a source as credible does not necessarily translate into personally endorsing its claim, raising questions about which factors ultimately influence personal decision-making beyond evaluation (Thomm & Bromme, 2016). Our findings suggest that although students generated a wide range of evaluation criteria, only a subset was activated when they actually sought to resolve expert disagreement. Moreover, personal aims and ideals, as well as ideals for conclusions, also mediated the transition from evaluation to resolution. Additionally, whereas researchers argued that different epistemic beliefs dimensions also linked to the integration of multiple texts (Bråten et al., 2011), our result extend this view by providing detailed and in process accounts of how students activate, weigh, coordinate, and apply a repertoire of epistemic ideals within specific contexts. These practices further shape document integration in diverse ways, ranging from incorporating corroborated information to ignoring conflicts and reorganizing information.

Our small sample size limits generalizability, but the verbal protocol analysis provides a fine-grained view of how personal and epistemic aims, ideals, and processes interact in real time. And we are analyzing an extra number of groups to develop a more robust quantitative analysis. Our findings point to the value of scaffolds that expose learners to a repertoire of ideals, create opportunities for explicit discussion of goals and ideals for conclusions, and support discussion about when and how to apply aims and ideals in combination, and which to prioritize in different contexts (Barzilai & Chinn, 2024).

References

Barzilai, S., & Chinn, C. A. (2024). The AIR and Apt-AIR Frameworks of epistemic performance and growth: Reflections on educational theory development. *Educational Psychology Review*, *36*(3), 91. https://doi.org/10.1007/s10648-024-09927-5

Barzilai, S., Mor-Hagani, S., Zohar, A. R., Shlomi-Elooz, T., & Ben-Yishai, R. (2020). Making sources visible: Promoting multiple document literacy with digital epistemic scaffolds. *Computers & Education*, *157*, 103980. https://doi.org/10.1016/j.compedu.2020.103980

Barzilai, S., & Strømsø, H. I. (2018). Individual differences in multiple document comprehension. In J. L. G. Braasch, I. Bråten, & M. T. McCrudden (Eds.), *Handbook of Multiple Source Use* (1st ed., pp. 99–116). Routledge. https://doi.org/10.4324/9781315627496-6

Barzilai, S., & Thomm, E. (2023, August 24). *Dealing with disagreement: coming to terms with the social nature of knowledge* [Video]. Invited keynote presentation at the 20th Biennial Conference of the European Association for Research on Learning and Instruction (EARLI), Thessaloniki, Greece. YouTube. https://www.youtube.com/watch?v=-Ri5TzrInx8

Barzilai, S., Thomm, E., & Shlomi-Elooz, T. (2020). Dealing with disagreement: The roles of topic familiarity and disagreement explanation in evaluation of conflicting expert claims and sources. *Learning and Instruction*, *69*, 101367. https://doi.org/10.1016/j.learninstruc.2020.101367

Barzilai, S., & Zohar, A. (2012). Epistemic thinking in action: Evaluating and integrating online sources. *Cognition and Instruction*, *30*(1), 39–85. https://doi.org/10.1080/07370008.2011.636495

Bewley, S. (2011). The NHS breast screening programme needs independent review. *BMJ*, *343*, d6894. https://doi.org/10.1136/bmj.d6894

Bråten, I., Britt, M. A., Strømsø, H. I., & Rouet, J.-F. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: Toward an integrated model. *Educational Psychologist*, *46*(1), 48–70. https://doi.org/10.1080/00461520.2011.538647

Bromme, R., & Goldman, S. R. (2014). The public's bounded understanding of science. *Educational Psychologist*, *49*(2), 59–69. https://doi.org/10.1080/00461520.2014.921572

Bromme, R., Thomm, E., & Wolf, V. (2015). From understanding to deference: Laypersons' and medical students' views on conflicts within medicine. *International Journal of Science Education, Part B*, *5*(1), 68–91. https://doi.org/10.1080/21548455.2013.849017

Broniatowski, D. A., Jamison, A. M., Qi, S., AlKulaib, L., Chen, T., Benton, A., Quinn, S. C., & Dredze, M. (2018). Weaponized health communication: Twitter bots and russian trolls amplify the vaccine debate. *American Journal of Public Health*, *108*(10), 1378–1384. https://doi.org/10.2105/AJPH.2018.304567

Chang, C. (2015). Motivated processing: How people perceive news covering novel or contradictory health research findings. *Science Communication*, *37*(5), 602–634. https://doi.org/10.1177/1075547015597914

Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *Journal of the Learning Sciences*, *6*(3), 271–315. https://doi.org/10.1207/s15327809jls0603_1

Chinn, C. A., & Duncan, R. G. (2018). What is the value of general knowledge of scientific reasoning? In F. Fischer, C. A. Chinn, K. Englemann, & J. Osborne (Eds.), *Scientific Reasoning and Argumentation: The Roles of Domain-Specific and Domain-General Knowledge* (pp. 77–101). Taylor and Francis. https://doi.org/10.4324/9780203731826

Chinn, C. A., & Rinehart, R. W. (2016). Epistemic cognition and philosophy: Developing a new framework for epistemic cognition. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 460-478). Routledge.

Clark, D., Nagler, R. H., & Niederdeppe, J. (2019). Confusion and nutritional backlash from news media exposure to contradictory information about carbohydrates and dietary fats. *Public Health Nutrition*, *22*(18), 3336–3348. https://doi.org/10.1017/S1368980019002866

Duncan, R. G., Chinn, C. A., & Barzilai, S. (2018). Grasp of evidence: Problematizing and expanding the next generation science standards' conceptualization of evidence. *Journal of Research in Science Teaching*, *55*(7), 907–937. https://doi.org/10.1002/tea.21468

Goldman, A. I. (2001). Experts: Which ones should you trust? *Philosophy and Phenomenological Research*, *63*(1), 85–110. https://doi.org/10.2307/3071090

Gottschling, S., Kammerer, Y., Thomm, E., & Gerjets, P. (2020). How laypersons consider differences in sources' trustworthiness and expertise in their regulation and resolution of scientific conflicts. *International Journal of Science Education, Part B*, *10*(4), 335–354. https://doi.org/10.1080/21548455.2020.1849856

Hellsten, I., Dawson, J., & Leydesdorff, L. (2010). Implicit media frames: Automated analysis of public debate on artificial sweeteners. *Public Understanding of Science*, *19*(5), 590–608. https://doi.org/10.1177/0963662509343136

Mochizuki, T., Chinn, C. A., Zimmerman, R. M., & Sekine, S. (2022). Promoting collaborative productive epistemic discourse for disagreement resolution among multiple documents: How epistemic scaffolds and epistemic scripts in CSCL worked. In A. Weinberger, W. Chen, D. Hernández-Leo, & B. Chen (Eds.), *Proceedings of the 15th International Conference on Computer-Supported Collaborative Learning (CSCL 2022)* (pp. 340–343). International Society of the Learning Sciences. https://repository.isls.org/handle/1/8302

Rouet, J.-F., & Britt, M. A. (2011). Relevance processes in multiple document comprehension. In *Text relevance and learning from text* (pp. 19–52). IAP Information Age Publishing.

Rouet, J.-F., Britt, M. A., & Durik, A. M. (2017). RESOLV: Readers' representation of reading contexts and tasks. *Educational Psychologist*, *52*(3), 200–215. https://doi.org/10.1080/00461520.2017.1329015

Stadtler, M., & Bromme, R. (2014). *The content–source integration model: A taxonomic description of how readers comprehend conflicting scientific information*. https://doi.org/10.7551/mitpress/9737.003.0023

Thomm, E., Barzilai, S., & Bromme, R. (2017). Why do experts disagree? The role of conflict topics and epistemic perspectives in conflict explanations. *Learning and Instruction*, *52*, 15–26. https://doi.org/10.1016/j.learninstruc.2017.03.008

Thomm, E., & Bromme, R. (2016). How source information shapes lay interpretations of science conflicts: Interplay between sourcing, conflict explanation, source evaluation, and claim evaluation. *Reading and Writing*, *29*(8), 1629–1652. https://doi.org/10.1007/s11145-016-9638-8

Thomm, E., Hentschke, J., & Bromme, R. (2015). The explaining conflicting scientific claims (ECSC) questionnaire: Measuring laypersons' explanations for conflicts in science. *Learning and Individual Differences*, *37*, 139–152. https://doi.org/10.1016/j.lindif.2014.12.001

Weinberger, A., Ertl, B., Fischer, F., & Mandl, H. (2005). Epistemic and social scripts in computer–supported collaborative learning. *Instructional Science*, *33*(1), 1–30. https://doi.org/10.1007/s11251-004-2322-4

Acknowledgements

This material is based upon work supported by the Grant-in-Aid for Scientific Research (B) (#JP23K25707) from Japan Society for the Promotion of Science, and the Telecommunication Advancement Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of these funding agencies.