Knowledge Construction and Uncertainty in Real World Argumentation: A Text Analysis Approach

Ha Nguyen
thicn@uci.edu
University of California, Irvine
Irvine, California, USA

William Young
wyoung@autodealerdata.com
Competitive Intelligence Solutions LLC.
Irvine, California, USA

ABSTRACT

Collaborative argumentation is key to promoting understanding of scientific issues. However, classroom structures may not always prepare students to engage in argumentation. To address this challenge, education researchers have examined the importance of social knowledge construction and managing uncertainty in group understanding. In this study, we explore these processes using data from /r/ChangeMyView, an online forum on Reddit where users present their opinions, engage others in critiquing ideas, and acknowledge when the discussion has modified their opinions. This unfacilitated environment can illuminate how argumentation evolves naturally towards refined opinions. We employ automated text analyses (LIWC) and discourse analyses to understand the features and discourse sequences of successful arguments. We find that argumentative threads are more likely to be successful if they focus on idea articulation, coherence, and semantic diversity. Findings highlight the role of uncertainty: threads with more certainty words are less likely to be successful. Furthermore, successful arguments are characterized by cycles of raising, managing, and reducing uncertainty, with more occurrences of evidence and idea incorporation. We discuss how learning environments can create norms for idea construction, coherence, and uncertainty, and the potential to provide adaptive prompts to maintain and reduce uncertainty when unproductive argumentative sequences are detected.

CCS CONCEPTS

• **Human-centered computing → Collaborative and social computing**.

KEYWORDS

argumentation, uncertainty, online communities, text analysis

ACM Reference Format:
Ha Nguyen and William Young. 2022. Knowledge Construction and Uncertainty in Real World Argumentation: A Text Analysis Approach. In *LAK22: 12th International Learning Analytics and Knowledge Conference (LAK22), March 21–25, 2022, Online, USA.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3506860.3506864

LAK22, March 21–25, 2022, Online, USA

ACM ISBN 978-1-4503-9573-1/22/03.
https://doi.org/10.1145/3506860.3506864

1 INTRODUCTION

Current science standards in the U.S. and international programs emphasize the importance of scientific literacy in workforce preparation and citizenry development [47, 59]. Individuals showcase scientific literacy through developing content knowledge and engaging in authentic inquiries in ways that are similar to science practitioners. A key aspect of scientific literacy is argumentation, where individuals propose, refine, and synthesize ideas [39]. Argumentation can enrich learning, particularly when it comes to discussing how scientific issues relate to everyday life [4, 23].

However, students often lack opportunities to engage in argumentation rigorously [50]. When faced with time constraints, teachers lean towards representing science concepts as facts and may not facilitate exploratory discussions among students [51]. As a result, students may only engage in shallow exchange and resist alternative perspectives in argumentation [22]. Collaborative argumentation can address these issues. This discourse form invites individuals to work together to surface ideas, resolve disagreement through elaboration, reflection, and reasoning, thereby developing deeper understanding of social and scientific issues [3, 11, 20, 50, 74]. Collaborative argumentation is successful when individuals contribute to *knowledge construction* to articulate and critique ideas [33]. In this process, individuals navigate *uncertainty* in group knowledge to build more coherent understanding over time [11, 13].

Researchers have examined collaborative argumentation in writing [44], classroom discussion [14], and teacher facilitation [12, 13]. Another research strand has attended to argumentation in out-of-school online discourse [23, 28, 29, 40]. Our research shares the premise that informal learning environments such as online discussions can foster argumentation skills [23]. Understanding how argumentation happens "in the wild", that is, in online communities without facilitation, may reveal insights into how to help students engage in productive discourse naturally [24, 31]. In addition, online environments provide rich data for large-scale, automated analyses. We leverage these affordances to explore how individuals engage in social knowledge construction and uncertainty, two processes critical to collaborative argumentation [11, 57, 67, 72].

In this work, we examine online, collaborative argumentation discourse from the /r/ChangeMyView (CMV) forum on Reddit (N = 8,167 messages, 1,954 discussion threads). Users submit their opinions to the forum to collect alternative perspectives from others. Participants explicitly acknowledge arguments that revise or expand their ideas by replying to the thread with a Delta (Δ). The Δ replies create natural categorizations of successful (marked by Δ) arguments that expand one's understanding versus unsuccessful ones. We apply these categories to understand factors that influence effective argumentation, using features that pertain to social knowledge construction and uncertainty [11, 30, 70, 74]. To illustrate the potential of large-scale analyses for this data source, we leverage LIWC (Linguistic Inquiry and Word Count; [52]) to develop the model features based on frequencies of linguistic markers.

Furthermore, we delve into how individuals navigate uncertainty throughout an argumentative thread. Prior work has not examined how individuals navigate uncertainty in argumentation beyond separate occurrences of discourse moves. Understanding the transitions between steps can assist future intelligent systems to provide the appropriate feedback for learners engaging in collaborative discourse. For this, we manually code and conduct process mining of the argumentative sequences for a subset of the threads. The following questions guide our analyses:

RQ1. What characterizes successful argumentation that results in refined understanding "in the wild"?

RQ2. How is uncertainty navigated in collaborative argumentation sequences that result in refined ideas, compared to unsuccessful sequences?

We find that on average, successful argumentative threads are associated with a higher focus on idea articulation (e.g., "I" or "my" ideas) and coherence (i.e., higher usage of conjunctions, longer responses). Our findings also illuminate the role of uncertainty in argumentation. Threads that included a higher number of certainty words were less likely to be successful. In addition, effective argumentation is characterized by cycles of raising uncertainty (e.g., asking questions to highlight discrepancies), maintaining uncertainty (e.g., offering alternative perspectives), and reducing uncertainty (e.g., synthesizing ideas). Meanwhile, unsuccessful threads often consist of only raising and maintaining uncertainty. This pattern indicates extensions of disagreement, without resolving the points of contention.

Our study contributes to the fields of learning analytics and educational sciences in two main ways. First, it illustrates the utility of leveraging public data mining [35] from unstructured online discussions to understand learning interactions. We show how linguistic features can be applied to explore theoretically-driven processes such as social knowledge construction and uncertainty. Second, findings on how individuals participate in collaborative argumentation serve as avenues for bridging in-school and out-of-school learning. For example, future systems to encourage argumentation can attend to social dynamics for idea articulation. These systems can provide feedback to reinforce cycles of raising, maintaining, and reducing uncertainty and intervene when the systems detect prolonged confusion.

2 BACKGROUND

2.1 Collaborative Argumentation from Learning Sciences Lenses

This study combines research in learning sciences and learning analytics to examine facets of collaborative argumentation. Collaborative argumentation involves individuals working together to resolve disagreement, instead of arguing out of competition [3]. Argumentation enriches learning through surfacing cognitive discrepancies, providing opportunities for reflection, and facilitating reasoning [3, 6, 7]. Learning sciences researchers have highlighted necessary structures for collaborative argumentation, such as social knowledge construction and uncertainty. Meanwhile, social learning analytics provides insights into methodologies such as text analyses to capture argumentation at scale.

We define **social knowledge construction** as the ways in which individuals engage in interactions with one another to develop, critique, and refine shared understanding [19, 25]. This definition builds on related concepts of social negotiation [11], knowledge building [30], and social construction of knowledge [17, 25] to center on social interactions and knowledge construction moves that develop group understanding [8, 64]. From a social perspective, individuals position themselves within a community committed to knowledge building [74], and thus transition fluidly between personal authority in idea articulation and group dialogues [5]. This transition can manifest in subtle cues, such as changes in pronouns from "we" (group responsibilities) to "you" (informational exchange) and "I" (individual responsibilities) to navigate roles and participation in group discourse [5, 69].

Individuals engage in an array of discourse moves to support knowledge construction [30]. They refer to shared ideas through consensus-building, from acknowledging a "not understanding" position to challenging, modifying, integrating, and accepting others' perspectives [13, 25, 72, 74]. Individuals construct increasingly coherent arguments and support ideas with reasoning [30, 74]. Furthermore, individuals build on previous utterances to construct shared knowledge [72]. Participants in argumentation not only refer to their own ideas; they also consider others' ideas and negotiate how these perspectives relate to the main argumentative threads [11, 64]. Idea negotiation influences individuals' language. It may converge with the group's language in homogeneous spaces but diverge in spaces with diverse opinions [2].

In social knowledge construction, **uncertainty** can trigger cognitive processes that promote decision-making and conceptual understanding [33, 42]. Learners deepen knowledge through recognizing and reconciling the dissonances between their understanding and their observations [58]. In the classroom, teachers and students problematize events and create moments of uncertainty to advance discussions [67]. For example, researchers have unpacked how teachers and students manage uncertainty in argumentation in three phases: raising, maintaining, and reducing uncertainty [11]. *Raising uncertainty* refers to strategies when individuals establish a need for deeper understanding (e.g., asking for clarification), identify weaknesses in argument, or recognize gaps in conceptual knowledge. This strategy utilizes discrepancies to challenge individuals' existing knowledge [33]. The group can *maintain uncertainty* through discussing solutions, soliciting evidence, and adding different perspectives. Finally, the group can connect new understanding to their current knowledge, revise misconceptions, and consolidate different perspectives into a coherent argument. This strategy is termed *reducing uncertainty*, as the groups reach resolutions on initial inconsistencies. In sum, the three stages of uncertainty management highlight discrepancies, maintain argumentation to prevent shallow exchange, and settle the discussion to form shared understanding [11, 57, 72].

2.2 Collaborative Argumentation from Social Learning Analytics

Formal learning structures often follow predetermined turn-taking norms (e.g., teacher poses a question, students respond, teachers ask follow-up questions). These norms shift in online environments that facilitate asynchronous exchange, anonymity, and open-ended discussions [24, 27, 31]. Social learning analytics account for these characteristics in online learning, and propose that individuals engage in learning exchange in interactions with networks of people, ideas, and resources [66]. Learning activities in online environments interface directly with others through messaging, following, networking and indirectly through activity traces, such as searching, posting, and rating messages [66].

Researchers have examined these learning traces in the wild to explore processes in a range of domains, such as persuasion and deception [1, 24, 29, 70, 73], response elicitation [16], argumentative sequences [28], and participation norms and community building [45, 68]. These studies often involve automated analyses of linguistic features to examine a large amount of discussion text. We thus build on the linguistic features and analytic approaches in prior work to conceptualize our frameworks of social knowledge construction and uncertainty.

2.2.1 Social knowledge construction. Building on our review of learning sciences research, we examine the social interactions and knowledge-building discourse of social knowledge construction [5, 30, 69]. Regarding social interactions, researchers have studied **pronoun usage** as reflective of a focus on social negotiation versus individuals' advancement of ideas [69, 70, 73]. Self reference with first-person pronouns (e.g., "I", "me", "my") may differ from group reference with first-person plurals (e.g, "we") or second-person ideas (e.g., "you") [45, 70]. Users who employ first-person pronouns in their original arguments are found to be more amenable to changing their opinions [70]. Participants in heterogeneous environments employ more second-person pronouns ("you"), whereas they are likely to refer to third-person plural pronouns ("we") in homogeneous spaces where everyone shares aligned opinions [2, 71].

To capture individuals' knowledge-building discourse, we draw from linguistic features in prior work: (1) how individuals provide reasoning to support ideas [74], (2) how arguments become increasingly coherent over time [11], and (3) how ideas from follow-up responses interface with prior arguments [2, 33, 64]. We outline these features below:

Reasoning, defined as the ability to cite evidence and explain reasoning in one's statement, increases the possibility that the group resolves discrepancies and advances more coherent understanding [11, 42, 74]. In prior work, reasoning statements have been captured through detection of causation words, such as "because", "effect", and "thus" [1].

To assess the specificity and coherence of the arguments, researchers have examined discourse markers and text length [2, 46, 70]. **Discourse markers**, such as conjunctions, negations, and exclusion words, require the speakers to be more precise in communication [46]. Thus, maintaining concise, articulate argumentation can characterize coherent knowledge exchange. In addition, **word count** has been identified as a strong predictor of effective threads in prior work [70]. Longer posts can be more explicit with more information, contributing to the argument's coherence [53, 54].

Finally, the **interplay of ideas** between subsequent posts and the original posting can reveal whether the unfolding discussion overlaps with or diverges from the original ideas [64, 70]. The relatedness of ideas between the original posters and subsequent users can describe how individuals refer to others' ideas to construct group understanding. For example, researchers have found that matching arguments to individuals' attitudes can enhance the text's persuasiveness [56]. In comparison, in one study with the same data context as ours (/r/ChangeMyView forum), researchers find that argumentative threads that are deemed more persuasive diverge more from the original posts than their counterparts [70]. Other researchers have particularly linked such idea divergence to interactions in heterogeneous communities [2].

2.2.2 Uncertainty. Participants communicate uncertainty in argumentation with linguistic markers [1, 49, 63]. Students in classrooms where teachers use more hedges words (e.g., "maybe", "guess", "perhaps") have more opportunities to evaluate each other's claims and evidence [49]. Researchers have also calculated occurrences of words that signify tentative arguments. **Tentative words** such as "may" or "perhaps" express lack of commitment, caution, or suggestion of an open position. In contrast, **certainty words** such as "actually" may indicate stronger commitments [1, 52].

2.3 Discourse Sequences to Navigate Uncertainty

Beyond linguistic features, the ways in which individuals structure their discourse in groups also supports the success of argumentation [28]. Models of collaborative argumentation point to fluidity in raising, maintaining, and reducing uncertainty [11]. Individuals highlight the discrepancies in data (i.e., raise uncertainty), discuss (i.e., maintain uncertainty), and resolve these issues (i.e., reduce uncertainty), before attending to the next areas of differences in the group's ideas (i.e., raise uncertainty). This model calls for examination of group exchange as sequences of interactions, rather than as occurrences of separate argumentative moves. Although prior studies provide the conceptual model for examining uncertainty navigation as sequences, few studies have applied sequence analyses methods to understand these patterns.

We turn to educational research for examples of sequence analyses of learning activities [10, 28, 36, 37, 43]. Researchers have mined sequences of discourse moves from verbal discussions, activity logs, and writing and creative artifacts to understand how groups co-develop their artifacts over time [43, 55], how individuals regulate efforts and contribute to collaborative learning activities [10, 48], and how individuals structure rhetorical moves in argumentative writing [36, 41]. A central premise of these studies is that sequence analyses reveal critical patterns in idea evolution and argument structures [37]. This body of work illuminates the insights that such analyses can reveal about how individuals share, communicate, and critique ideas in collaborative exchange. Attending to sequential interactions can help detect when a participant's opinion has changed and the discourse moves leading up to that moment.

In sum, online communities provide rich environments to study how individuals engage in collaborative argumentation. Text analysis based on linguistic features offers an alternative approach to analyzing online interactions and conceptualizing constructs such as knowledge construction and uncertainty. In our research, we combine automated text analyses with qualitative insights to (1) illuminate how collaborative argumentation can be examined at scale, and (2) explore how argumentative sequences unfold over time. We examine the characteristics of successful argumentative threads and the sequences of argumentative moves to manage uncertainty, an important facet to guide the discussion.

3 METHODOLOGIES

3.1 Data Source

The data for this study came from the /r/ChangeMyView (CMV) forum on the Reddit platform. The CMV forum encourages individuals to post original opinions and engage in conversations with other users to change their position. Researchers have employed CMV data to examine argumentation [28, 45, 70]. The arguments are generally high-quality, since they are monitored to flag abusive comments or posts with no substantial arguments. Each discussion in CMV centers around the opinions posted by the original posters (OPs) and forms a discussion tree of comments and replies to the comments. Both the OPs and participants in a discussion thread can award Deltas (Δ) to any argument that convinces them to change their views or broadens their original opinions. Within the same thread, Deltas have internal consistency as the posters justify how a comment deserves an award.

In this work, we explored features in social knowledge construction and uncertainty that differentiated a successful argumentation thread (that was awarded a Δ) from unsuccessful ones. We used the Reddit API to sample 554 original posts on CMV created in 2021, retrieved all comment threads in response to these posts that received a Δ and randomly retrieved a roughly equivalent number of threads without Δ that responded to the same posts (1,954 discussion threads in total, 914 Δ, 1,041 non-Δ threads). Our goal was to identify arguments that result in refined opinions, as acknowledged by the users. The Deltas explicitly marked when users acknowledged having changed their opinions, and were thus sufficient for our purpose. We defined a thread to include the original posts, a comment, and its associated replies (Figure 1). To make sure that posters were engaged in the comments, we selected threads with a depth of at least two. This means that the threads involved the original post, a comment to the post, and at least one reply to the comment. Figure 1 provides examples of the Δ and non-Δ threads under the same post.

3.2 Analytical Procedures

3.2.1 RQ1: Characteristics of Successful Argumentation. To predict successful argumentative threads, we selected features based on frameworks in social knowledge construction and uncertainty stemming from learning contexts [11, 30, 70, 74]. Table 1 provides an overview of the features. Several features, including discourse markers, reasoning, and uncertainty, were derived from LIWC [52]. LIWC is an automated text classification program that groups words into categories to represent components of emotion and cognition, and calculates the frequencies that words in the text belong to these categories. LIWC has been applied to study the development of online learning communities [1, 24].

To measure the interplay of ideas between comments and the original posts, we used Universal Sentence Encoder (USE; [9]) to calculate the semantic similarity between posts. USE turned the CMV posts and comments into high-dimensional vectors that can be used for classification, clustering, and other natural language tasks. Two sentences that were closer in meaning would have a smaller Cosine distance between their respective vectors, and consequently a higher USE similarity score. For example, when comparing the sentence "It's important to distinguish between prevention and mitigation" to "I want cake", USE returned .03 (in a possible range of 0-1), whereas comparing the first sentence to a third sentence ("For both prevention and mitigation, things we can do feels minimal.") returned a score of .76. The validation section details how we compared results from USE with those from a human coder.

We performed a logistic regression model to understand factors that influenced Δ (successful) threads. We standardized all features to account for unit variance. Because the data were collected for individual posts but the prediction was for the threads (i.e., group-level outcome), we calculated the average of the features from Table 1 within threads. This calculation introduced heteroscedasticity, meaning that the residual errors did not have the same variance across observations. Thus, we used White's heteroscedasticity adjustment to calculate robust standard errors. This strategy has been found to achieve more power for analyses of group-level outcomes [21].

We note that our analytical approach focused on interpreting the features related to our theoretical constructs, as opposed to devising the most accurate prediction models from a broad feature set (for examples of the latter goal, see [1, 45, 70]). We chose logistic regression models for their interpretability affordances.

3.2.2 RQ2: Sequences of Navigating Uncertainty in Argumentation. Given the importance of managing uncertainty in collaborative argumentation [11, 33, 57, 72], for RQ2, we examined the process that groups raised, maintained, and reduced uncertainty. For this, we conducted manual coding for a randomly selected set of 200 argumentative threads (100 Δ, 100 non-Δ) from the larger sample in RQ1.

We developed a coding scheme based on prior frameworks with classroom argumentation [11, 33] to identify discourse moves for navigating uncertainty. The coding scheme included three high-level categories (raising, maintaining, reducing uncertainty) and sub-codes under each category. Raising uncertainty involves clarifying acts and identifying discrepancies in the group's understanding. Maintaining uncertainty refers to providing claims or evidence to strengthen the arguments. Finally, participants can reduce uncertainty by connecting new ideas or synthesizing perspectives into more coherent arguments. We added an Off-task category for comments that were unrelated to the posts they were responding to. Table 2 provides examples of the coding scheme. Each comment (excluding the original post) received one code and one sub-code for the primary discourse. For example, a comment that included one clarifying question but multiple evidence statements would receive "Maintain uncertainty" for the main code and "Evidence" for

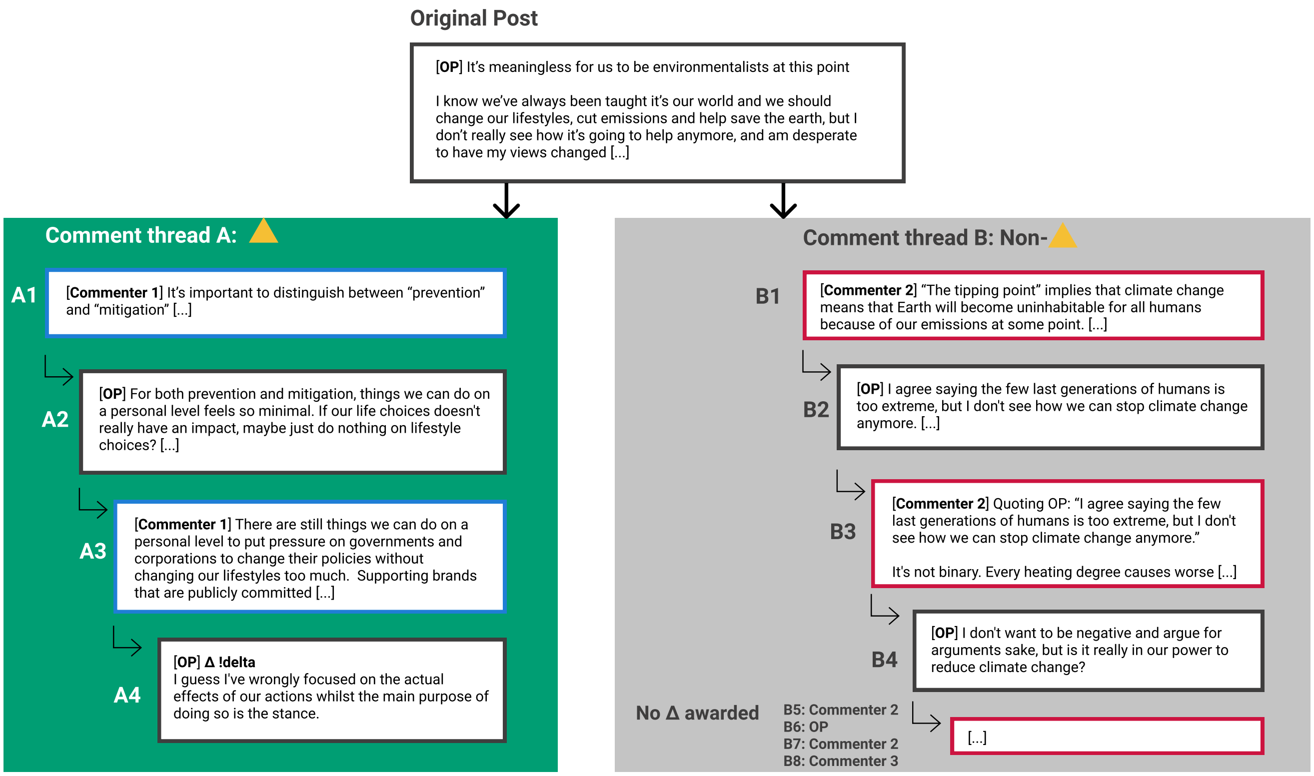

Figure 1: Example discussions on /r/ChangeMyView forum for Δ and non-Δ threads under the same original post. Text box's colors indicate different users. The poster explicitly acknowledged opinion refinement in A4.

Table 1: Features to Predict Successful Argumentation, Drawing from Learning Settings

Feature name	Definitions/Examples	Hypothesized relations with Δ thread
SOCIAL KNOWLEDGE CONSTRUCTION		
Pronoun: I	I, me, mine	+ [2, 5, 45, 69, 70]
Pronoun: you	you, your, thou	+ [69]
Pronoun: we	we, us, our	- [2, 70, 71]
Similarity	Semantic similarity with original posts	- [70]
Word counts	Counts of words in posts	+ [54, 70]
Discourse markers: Conjunctions	and, but, whereas	+ [1, 46]
Discourse markers: Negations	no, not, never	+ [1, 46]
Reasoning	hence, thus, because	+ [30, 53, 54]
UNCERTAINTY		
Tentative	maybe, perhaps	+ [11, 49, 63]
Certainty	always, never	- [1]

Notes: + = positive (an increase in this variable is associated with the increased likelihood of a Δ thread)
- = negative relation (an increase in this variable is associated with the decreased likelihood of a Δ thread)

the sub-code. The sequence for an argumentative thread can take the form [Raise uncertainty -> Maintain uncertainty -> Maintain uncertainty] following the original post. To establish inter-rater reliability, an undergraduate research assistant not involved in the study's conceptualization was trained on the coding scheme. The first author and the assistant coded 20% of the data separately, and achieved acceptable agreement (κ = .79). The first author coded the rest of the data.

We conducted three types of analyses on the coded data to examine discourse that facilitated uncertainty. All analyses compared

Table 2: Examples of uncertainty discourse moves

Discourse move	Definition	Example
Raising uncertainty		
Clarify	Establish a need for understanding	Can you clarify?
Identify discrepancies	Identify discrepancies or weaknesses in argument	Your argument doesn't follow. What if I make 40k? ...
Maintaining uncertainty		
Claim	Provide claims to strengthen arguments	... but my argument is for companies that continue to let employees work from home.
Evidence	Provide evidence to add or assimilate perspectives	Right now, my county has about 370k people and is getting about 300 new cases per day.
Reducing uncertainty		
Synthesize ideas	Consolidate arguments into a coherent proposition	In my defense, **my post** is mostly in reference to automoderation that focuses specifically on detecting words and phrases. [...] , or **like you said**, spam prevention are more reasonable.
Incorporate new ideas	Connect new information to existing knowledge	You do make a very good point about post history !delta.
Off-task	Discussion not connected to original posts	Irrelevant, but where did you get your degree?

two groups: successful (with Δ) and unsuccessful threads (no Δ). First, we created **process maps** of the high-level codes for raising, maintaining, and reducing uncertainty to examine how individuals in the two groups navigated argumentation over time. We transformed the data into event log format, where each event represented an entry (an original post or comment by a user) and an event sequence represented a discussion thread. The process maps served as an exploratory analysis to show the relative frequencies of transitioning from one state (e.g., raising uncertainty) to another (e.g, maintaining uncertainty) or engagement in self-loops (i.e., showing the same uncertainty moves several times). The darker colors in the process maps indicated higher frequencies. Second, we took a closer look at the sub-codes (e.g., clarify, evidence, claim) to gain a deeper understanding of differences in strategies between groups. We used **Mann-Whitney tests** to compare the frequencies of discourse moves among threads with and without Δ. Third, we calculated the **transition probabilities** of sub-codes. The transition probabilities showed the likelihood that a discourse move (e.g., claim) shifted to another (e.g., evidence) in one step. We used Pearson's chi-squared tests of independence to determine if there existed significant differences in the transition probabilities between groups.

The analyses used LIWC [52] to process the linguistic features, TensorFlow Hub's Universal Sentence Encoder [9] to calculate semantic similarity, and R to calculate the logistic model, robust standard errors [60], process mapping [32], and transition probabilities [65].

3.3 Validation

We attempted to validate the relative ranking from Universal Sentence Encoders's (USE; [9]) semantic similarity output with assessments from a human coder. We pulled a random sample of 5% of the data (30 threads, 105 messages). A research assistant blinded to the USE results rated pairs of original post - comments on a scale of 1 to 5 (1 being very different in content, and 5 being very similar). This rating scheme grouped the post-comment pairs into five groups, and we compared the group distribution in the human coder's rating with the USE output (also divided into five groups based on the USE similarity output, 0-.2, .2-.4, .4-.6, .6-.8, .8-1). 86% of the pairs were put in the same groups by the human coder and USE, suggesting substantial agreement.

Table 3: Predictors of Successful Argumentative Threads

Feature	OR	Robust SE	z	p
Knowledge construction				
Pronoun: I	1.28	.07	4.33	***
Pronoun: you	.77	.05	-4.39	***
Pronoun: we	.87	.08	-1.63	.10
Semantic similarity	.59	.04	-8.32	***
Word counts	5.26	.92	9.48	***
Conjunctions	1.16	.07	2.61	.01**
Negations	.90	.05	-1.96	.05
Reasoning	1.14	.05	1.89	.06
Uncertainty				
Tentative	.99	.05	-.17	.86
Certainty	.89	.05	-2.20	.03*

Notes: *: $p < .05$; **: $p < .01$; ***: $p < .001$

4 RESULTS

4.1 RQ1: Characteristics of Successful Argumentation

We first explored the features of successful argumentation threads in an online community. Identifying the factors that influenced argumentation in these contexts can inform learning facilitation. For this, we ran a logistic regression model using features of social knowledge construction (pronoun usage, interplay of ideas, discourse markers, reasoning) and uncertainty (usage of words that suggest a tentative or certain propositions). Table 3 reports the results. Below we outline the features that contributed to the model in meaningful ways (Wald test $p < .05$).

We found that threads with a higher average proportions of individual, first-person pronouns (e.g., "I") were more likely to be successful (OR = 1.28, SE = .07, z = 4.33, $p < .001$). These findings indicate that individuals in effective threads appeared to focus on idea articulation overall. Findings echo prior analyses on pronouns in heterogeneous space where individuals hold different opinions [2]. Meanwhile, a higher average frequency of second-person pronouns (e.g., "you") was associated with a lower likelihood of success (OR = .77, SE = .05, z = -4.39, $p < .001$). This finding differs from for our hypotheses that individuals engaging in the construction and critique of ideas would be more likely to reference each other's ideas through second-person pronouns [69]. It is possible that critiques of ideas that focus on others' responses can gear groups towards polarization rather than consensus [26].

Successful threads also appeared to contain more details on average. They were generally associated with a greater number of words (OR = 5.26, SE = .92, z = 9.48, $p < .001$) and a higher usage of conjunctions to connect sentences (OR = 1.16, SE = .07, z = 2.61, $p = .01$). Conjunctions are viewed as markers of thought complexity, as they help to connect strands of independent ideas to build more coherent discourse [61]. Interestingly, effective threads appeared to have less semantic similarity between follow-up replies and the original posts, compared to ineffective threads (OR = .59, SE = .04, z = -8.32, $p < .001$). This means that successful arguments diverged in meaning from the original posts.

Finally, we found that threads with more certainty words on average were less likely to be successful (OR = .89, SE = .05, z = -2.20, $p = .03$). Findings align with learning sciences research that uncertainty is critical to navigating discussions [49]. Problematizing discrepancies in understanding and assuming a less certain stance can help to advance the discussion and deepen participants' understanding [42, 58].

4.2 RQ2: Sequences of Navigating Uncertainty

To understand how to best facilitate groups' navigation of uncertain discourse, we examined the process maps (Figure 2) of raising, maintaining, and reducing uncertainty. In the figure, the percentages indicate the relative frequencies of transitioning between states and darker colors indicate higher frequencies. Descriptively, successful threads appeared to contain more frequent links between raising uncertainty (e.g., clarifying and asking for evidence), maintaining uncertainty (e.g., introducing discrepancies or providing evidence), and reducing uncertainty (e.g., incorporating new ideas or synthesizing arguments). Meanwhile, unsuccessful threads consisted of more cycles of raising and maintaining uncertainty and less frequent transitions between these phases and reducing uncertainty. For example, Figure 2 shows that once the discussion entered maintaining uncertainty stage, the transition frequency between maintaining and reducing uncertainty in Δ threads was 29.76%, compared to 2.88% in non-Δ threads. These patterns may indicate extended discrepancies without consensus.

We then examined the sub-codes for uncertainty in two ways: as frequencies of occurrences and transition probabilities. These analyses provided a more nuanced understanding of the strategies that individuals employed in co-constructing arguments. Mann-Whitney tests revealed that evidence and incorporation of new ideas occurred with higher frequencies in successful threads than unsuccessful ones (Evidence: M Δ = .62, SD = .75, M non-Δ = .20, SD = .48, $p = .01$; Incorporating new ideas: M Δ = .88, SD = .52, M non-Δ = .17, SD = .38, $p < .001$). These results highlight the importance of providing evidence and recognizing new ideas to guide groups towards consensus [11, 30, 50]. Table 4 lists the descriptive statistics and results from the Mann-Whitney tests.

The transition probability matrices (Figure 3) provide an additional lens to examine how groups moved from raising and maintaining uncertainty to idea synthesis in successful threads. At first glance, the darker-colored boxes in the Δ threads (left panel, Figure 3) illustrate that successful threads showed higher probabilities of transitioning from other discourse moves to incorporating ideas, synthesizing ideas, and identifying discrepancies. Pearson's chi-squared tests revealed similar patterns: the transition probabilities between other discourse moves and incorporating new ideas differed between successful and unsuccessful threads (χ^2 = 14.11, df = 5, $p = .01$). In particular, post-hoc tests suggested differences in the transition probabilities from identifying discrepancies, making claims, and providing evidence to incorporating new ideas between the groups (χ^2 = 17.19, 10.31, 10.31, respectively; df = 1, $p < .01$).

To interpret these quantitative findings, we qualitatively reviewed the unsuccessful threads that contained evidence, claim-making, or idea synthesis but did not lead to incorporating new ideas. In those unsuccessful threads, we observed that the discourse moves were either followed by more uncertainty (e.g., follow-up responses continued to identify discrepancies) or abandoned by the participants, thus never resolving the focal points of contention. Comment thread B in Figure 1 serves as an example for raising uncertainty (B4), without resolving disagreements.

5 DISCUSSION

5.1 Role of Social Knowledge Construction in Collaborative Argumentation

In this study, we explored the features that predicted successful argumentation—one that resulted in refined or expanded opinions. We based our feature selection on learning sciences frameworks in social knowledge construction and uncertainty, many of which were grounded in learning settings [11, 42, 74]. Regarding social dynamics, findings extend prior work that focuses only on pronouns in the original posts [70] to pronoun usage across the whole argumentative threads. While examining the original posts reveals insights into the posters' receptivity to alternative perspectives, understanding pronoun usage throughout an argumentative threads can illustrate productive social exchange norms.

We found that successful threads tended to feature more articulation of individuals' ideas. For example, a higher usage of first-person pronouns ("I") may indicate open-mindedness and more involved participation to support effective arguments [15, 70]. At the same time, we did not find a significant relation between usage of first-person plural pronouns ("we") and successful arguments. Prior research has attributed this finding to the discussion norms in the /r/ChangeMyView forum [70]. For example, Tan and colleagues

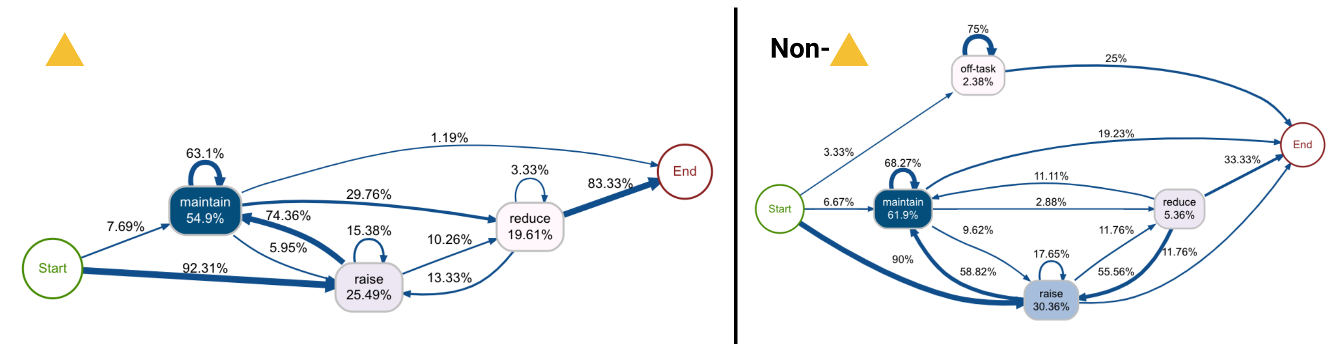

Figure 2: Process maps to show distribution of raising, maintaining, and reducing uncertainty in Δ and non-Δ threads

Table 4: Occurrences of Uncertainty Sub-codes within an Argumentative Thread

Discourse move	M Δ	SD Δ	M non-Δ	SD non-Δ	W	p
Raising uncertainty						
Clarify	.42	.70	.57	.77	430	.45
Identify discrepancies	1.73	1.37	2.07	1.76	423	.58
Maintaining uncertainty						
Claim	1.85	1.26	2.2	1.4	453.5	.29
Evidence	.62	.75	.20	.48	261	.01**
Reducing uncertainty						
Synthesize ideas	.38	.57	.27	.45	355	.48
Incorporate new ideas	.88	.52	.17	.38	135	***
Off-task	0	0	.13	.73	403	.38

Notes: N = 200 (100 Δ, 100 non-Δ). *: p < .05; **: p < .01; ***: p < .001

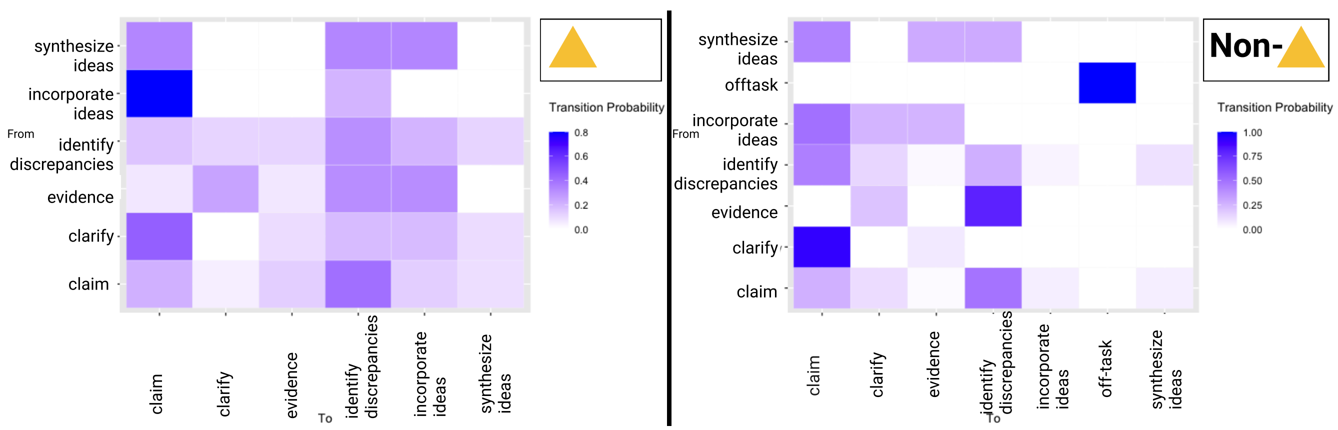

Figure 3: Transition probabilities for raising, maintaining, and reducing uncertainty. Darker colors represent larger probabilities.

[70] hypothesize that using the first-person plural pronouns may dilute the group responsibilities for navigating the opinions under discussion. Future research can examine pronoun usage in other contexts of collaborative argumentation to understand the shifts between first and second-person pronouns, as well as between individual and group identities.

Our findings also point to the importance of reasoning and idea coherence. Consistent with our hypotheses, we found that longer arguments and more frequent uses of conjunctions were positively correlated with the likelihood of an effective argument. These features imply more sophisticated thought processes [61]. Researchers have found similar links between greater word use, complex reasoning, and communication in online spaces with heterogeneous opinions, compared to homogeneous space [2]. Findings that higher semantic similarity in subsequent posts was associated with lower odds of effective threats stands in contrast with expectations for overlap of argumentative content in growing communities [56]. Instead, our findings point towards idea divergence in heterogeneous space and the importance of surfacing novel claims, alternative perspectives, and evidence in co-constructing group arguments [2, 11]. These findings have implications for facilitating argumentation in learning settings, for example, to surface alternative perspectives.

Overall, findings highlight exchange styles where individuals focus on argument articulation with more sophistication, compared to unsuccessful argumentative threads. Individuals in learning contexts can facilitate collaborative argumentation through articulating ideas, with a focus on reasoning and diverse contributions to group perspectives.

5.2 Managing Uncertainty in Collaborative Argumentation

In general, we found that argumentative threads that used more certainty words (e.g., "always", "never") were less likely to be successful. This finding supports learning sciences frameworks, which propose that taking a tentative stance can facilitate idea articulation and evaluation [11, 33, 49]. While uncertainty in posing opinions can be treated as deficient in classroom contexts [38], providing space for alternative discussion and collaborative argumentation can be conducive to developing scientific knowledge [50].

Analyses in RQ2 further illustrate instances where uncertainty can be facilitated. A large number of unproductive threads included cycles of raising and maintaining uncertainty, without synthesizing ideas. In contrast, we observed more transitions from raising and maintaining to reducing uncertainty within the productive threads. Our findings provide empirical evidence to support the uncertainty stages proposed in prior work [11, 42], and suggest parallel patterns between classroom observations and online communities. Furthermore, findings echo the concept of productive failure in science and mathematics education [34]. Facilitation strategies to promote productive failure introduce learners to active moments of uncertainty that prompt them to seek out more coherent arguments. These strategies differ from direct scaffolds and feedback to guide learners towards the correct solution. Through this process, learners generate diverse ideas and explanations and build skills to address more complex problems in subsequent tasks [34].

The transition probability matrices provide some guidance on how uncertainty can be managed successfully. Successful threads were more likely to show transitions from other discourse moves to incorporation of new ideas. In particular, individuals were more likely to move from giving evidence, making claims, and identifying discrepancies in previous ideas to recognition of new perspectives. These patterns suggest that argumentation can thrive in environments where participants actively co-construct ideas [33]. The argumentative space must encourage participants to identify gaps in their understanding, provide claims and evidence to support their interpretations, and most importantly, find ways to connect new ideas to the ongoing discourse [25]. Dialogues and negotiation play crucial roles in reducing discrepancies and reconciling alternative perspectives in collaborative argumentation.

Our results illustrate how uncertainty can be incorporated into argumentative threads in learning settings to help participants transition from discourse moves that raise and maintain uncertainty to resolving discrepancies. For instance, teachers can equip students with epistemic resources to focus on all stages of managing uncertainty [11]. These resources come in the form of teachers' framing of the activities, modeling of argumentative discourse, and discussion prompts for student discussion [12, 13, 44]. Similarly, intelligent learning systems or activity dashboards can keep track of the stages in which groups are navigating uncertainty, and provide the appropriate scaffolds if the systems detect unproductive extension of raising-managing uncertainty without idea synthesis or incorporation of new ideas.

6 LIMITATIONS

The limitations in the current research can guide future inquiries. First, we limited the features to knowledge construction and uncertainty based on our theoretical frameworks. Future research can thus consider a broader set of features, such as emotions, authority, and power dynamics, to gain better insights into characteristics of collaborative argumentation.

Second, our analyses did not differentiate arguments by topics on the /r/ChangeMyView forum. It is possible that there exist differences in the strategies and argumentation sequences across social, political, and scientific discussion topics. Thus, researchers in future work can compare linguistic strategies across topics. A related question for future research is how findings from this study generalize to other argumentation environments beyond the /r/ChangeMyView forum. These environments may involve different social interactions and argumentative strategies.

Third, to understand the sequences of argumentation, we only conducted manual coding for a small subset of the data. Future iterations can leverage linguistic features (e.g., sequences of discourse connectors, questions, and causal relationships) and experiment with classification models to automate the detection of uncertainty phases. Such an understanding builds towards creating systems that detect unproductive cycles of uncertainty management and provide the appropriate nudges. Collaborative exchange presents robust opportunities to examine such adaptive support [18, 62].

7 CONCLUSION

In this study, we illustrate the utility of mining public online discourse to examine collaborative argumentation. We find that a higher focus on articulation of individuals' ideas was likely associated with open-mindedness and a higher likelihood of successful argumentative threads. In general, threads that contained more indicators of sophisticated thought processes (i.e., higher usage of conjugations and higher word count) were more likely to be effective. Additionally, semantic divergence from the original opinions was predictive of successful threads. This finding highlights the need to encourage idea diversity in collaborative argumentation.

Results also reveal the link between uncertainty and successful arguments. Threads with a higher average use of certainty words were less likely to result in refined opinions. More importantly, successful threads more often contained cycles of raising, maintaining, and reducing uncertainty in group understanding. We unpacked how participants sequenced their discourse moves, and found that participants in successful threads were more likely to transition from other discourse moves to incorporation of new ideas. These findings have implications for other learning contexts. Findings provide support for intentionally creating spaces of managing and reducing uncertainty to elevate group understanding. They also highlight the importance of intervening with the right resources to avoid confusion and drawn-out disagreement.

Overall, this study demonstrates how to leverage linguistic features to explore unstructured argumentation processes. We draw parallels between social knowledge construction in classroom learning and unfacilitated online communities, thus showing the potential of online environments to inform learning design. Our findings build towards integrating social elements, such as facilitation of idea articulation, semantic diversity, coherence, and moments of uncertainty, to make argumentation more engaging in learning contexts.

ACKNOWLEDGMENTS

We would like to thank the research assistant in helping us refine the coding scheme and establish inter-rater reliability. We also thank the reviewers for their helpful feedback to strengthen the manuscript.

REFERENCES

[1] Aseel Addawood, Adam Badawy, Kristina Lerman, and Emilio Ferrara. 2019. Linguistic Cues to Deception: Identifying Political Trolls on Social Media. *Proceedings of the International AAAI Conference on Web and Social Media* 13 (7 2019), 15–25. https://ojs.aaai.org/index.php/ICWSM/article/view/3205

[2] Jisun An, Haewoon Kwak, Oliver Posegga, and Andreas Jungherr. 2019. Political Discussions in Homogeneous and Cross-Cutting Communication Spaces. *Proceedings of the International AAAI Conference on Web and Social Media* 13 (7 2019), 68–79. https://ojs.aaai.org/index.php/ICWSM/article/view/3210

[3] J E B Andriessen. 2006. Arguing to learn. (2006).

[4] Sasha A. Barab, Troy D. Sadler, Conan Heiselt, Daniel Hickey, and Steven Zuiker. 2006. Relating Narrative, Inquiry, and Inscriptions: Supporting Consequential Play. *Journal of Science Education and Technology 2006 16:1* 16, 1 (11 2006), 59–82. https://doi.org/10.1007/S10956-006-9033-3

[5] Leema K. Berland and David Hammer. 2012. Framing for scientific argumentation. *Journal of Research in Science Teaching* 49, 1 (1 2012), 68–94. https://doi.org/10.1002/TEA.20446

[6] John D Bransford, Ann L Brown, and Rodney R Cocking. 2000. *How people learn*. Vol. 11. Washington, DC: National academy press.

[7] Leah A. Bricker and Philip Bell. 2008. Conceptualizations of argumentation from science studies and the learning sciences and their implications for the practices of science education. *Science Education* 92, 3 (5 2008), 473–498. https://doi.org/10.1002/SCE.20278

[8] Zoë E. Buck, Hee-Sun Lee, and Joanna Flores. 2014. I Am Sure There May Be a Planet There: Student articulation of uncertainty in argumentation tasks. *http://dx.doi.org/10.1080/09500693.2014.924641* 36, 14 (2014), 2391–2420. https://doi.org/10.1080/09500693.2014.924641

[9] Daniel Cer, Yinfei Yang, Sheng-yi Kong, Nan Hua, Nicole Limtiaco, Rhomni St John, Noah Constant, Mario Guajardo-Cespedes, Steve Yuan, and Chris Tar. 2018. Universal sentence encoder for English. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*. 169–174.

[10] Bodong Chen, Monica Resendes, Ching Sing Chai, and Huang-Yao Hong. 2017. Two tales of time: uncovering the significance of sequential patterns among contribution types in knowledge-building discourse. *Interactive Learning Environments* 25, 2 (2 2017), 162–175. https://doi.org/10.1080/10494820.2016.1276081

[11] Ying-Chih Chen, Matthew J. Benus, and Jaclyn Hernandez. 2019. Managing uncertainty in scientific argumentation. *Science Education* 103, 5 (9 2019), 1235–1276. https://doi.org/10.1002/SCE.21527

[12] Ying-Chih Chen, Brian Hand, and Lori Norton-Meier. 2017. Teacher roles of questioning in early elementary science classrooms: A framework promoting student cognitive complexities in argumentation. *Research in Science Education* 47, 2 (2017), 373–405.

[13] Christine Chin and Jonathan Osborne. 2010. Supporting Argumentation Through Students' Questions: Case Studies in Science Classrooms. *http://dx.doi.org/10.1080/10508400903530036* 19, 2 (4 2010), 230–284. https://doi.org/10.1080/10508400903530036

[14] Andri Christodoulou and Jonathan Osborne. 2014. The science classroom as a site of epistemic talk: A case study of a teacher's attempts to teach science based on argument. *Journal of Research in Science Teaching* 51, 10 (12 2014), 1275–1300. https://doi.org/10.1002/TEA.21166

[15] Geoffrey L Cohen, Joshua Aronson, and Claude M Steele. 2000. When beliefs yield to evidence: Reducing biased evaluation by affirming the self. *Personality and social psychology bulletin* 26, 9 (2000), 1151–1164.

[16] . Danish, Yogesh Dahiya, and Partha Talukdar. 2016. Discovering Response-Eliciting Factors in Social Question Answering : A Reddit Inspired Study. *Tenth International AAAI Conference on Web and Social Media* (3 2016). https://www.aaai.org/ocs/index.php/ICWSM/ICWSM16/paper/view/13152

[17] Bram De Wever, Hilde Van Keer, Tammy Schellens, and Martin Valcke. 2010. Roles as a structuring tool in online discussion groups: The differential impact of different roles on social knowledge construction. *Computers in Human Behavior* 26, 4 (7 2010), 516–523. https://doi.org/10.1016/J.CHB.2009.08.008

[18] Dejana Diziol, Erin Walker, Nikol Rummel, and Kenneth R. Koedinger. 2010. Using intelligent tutor technology to implement adaptive support for student collaboration. , 89–102 pages. https://doi.org/10.1007/s10648-009-9116-9

[19] Maria Evagorou and Jonathan Osborne. 2013. Exploring young students' collaborative argumentation within a socioscientific issue. *Journal of Research in Science Teaching* 50, 2 (2 2013), 209–237. https://doi.org/10.1002/TEA.21076

[20] Mark Felton, Amanda Crowell, and Tina Liu. 2015. Arguing to Agree: Mitigating My-Side Bias Through Consensus-Seeking Dialogue. *http://dx.doi.org/10.1177/0741088315590788* 32, 3 (7 2015), 317–331. https://doi.org/10.1177/0741088315590788

[21] Lynn Foster-Johnson and Jeffrey D. Kromrey. 2018. Predicting group-level outcome variables: An empirical comparison of analysis strategies. *Behavior Research Methods 2018 50:6* 50, 6 (3 2018), 2461–2479. https://doi.org/10.3758/S13428-018-1025-8

[22] Merce Garcia-Mila, Sandra Gilabert, Sibel Erduran, and Mark Felton. 2013. The Effect of Argumentative Task Goal on the Quality of Argumentative Discourse. *Science Education* 97, 4 (7 2013), 497–523. https://doi.org/10.1002/SCE.21057

[23] Christine Greenhow, Thor Gibbins, and Melissa M. Menzer. 2015. Re-thinking scientific literacy out-of-school: Arguing science issues in a niche Facebook application. *Computers in Human Behavior* 53 (12 2015), 593–604. https://doi.org/10.1016/J.CHB.2015.06.031

[24] Anatoliy Gruzd, Priya Kumar, Deena Abul-Fottouh, and Caroline Haythornthwaite. 2020. Coding and Classifying Knowledge Exchange on Social Media: a Comparative Analysis of the #Twitterstorians and AskHistorians Communities. *Computer Supported Cooperative Work (CSCW) 2020 29:6* 29, 6 (6 2020), 629–656. https://doi.org/10.1007/S10606-020-09376-Y

[25] Charlotte N. Gunawardena, Constance A. Lowe, and Terry Anderson. 1998. Analysis of a Global Online Debate and the Development of an Interaction Analysis Model for Examining Social Construction of Knowledge in Computer Conferencing:. *http://dx.doi.org/10.2190/7MQV-X9UJ-C7Q3-NRAG* 17, 4 (2 1998), 397–431. https://doi.org/10.2190/7MQV-X9UJ-C7Q3-NRAG

[26] Ahmed Hassan, Vahed Qazvinian, and Dragomir Radev. 2010. What's with the attitude? identifying sentences with attitude in online discussions. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*. 1245–1255.

[27] Caroline Haythornthwaite and Richard Andrews. 2011. *E-learning theory and practice*. Sage Publications.

[28] Christopher Hidey and Kathleen McKeown. 2018. Persuasive Influence Detection: The Role of Argument Sequencing. *Proceedings of the AAAI Conference on Artificial Intelligence* 32, 1 (4 2018). https://ojs.aaai.org/index.php/AAAI/article/view/12003

[29] Christopher Hidey, Elena Musi, Alyssa Hwang, Smaranda Muresan, and Kathleen McKeown. 2017. Analyzing the semantic types of claims and premises in an online persuasive forum. In *Proceedings of the 4th Workshop on Argument Mining*. 11–21.

[30] Cindy E. Hmelo-Silver and Howard S. Barrows. 2008. Facilitating Collaborative Knowledge Building. *Cognition and Instruction* 26, 1 (1 2008), 48–94. https://doi.org/10.1080/07370000701798495

[31] Christopher Hoadley and Yael Kali. 2019. Five Waves of Conceptualizing Knowledge and Learning for Our Future in a Networked Society. *Learning In a Networked Society* 17 (2019), 1–21. https://doi.org/10.1007/978-3-030-14610-8{_}1

[32] Gert Janssenswillen and Benoît Depaire. 2017. bupaR: Business Process Analysis in R.. In *BPM (Demos)*.

[33] Michelle E. Jordan and Reuben R. McDaniel Jr. 2014. Managing Uncertainty During Collaborative Problem Solving in Elementary School Teams: The Role of Peer Influence in Robotics Engineering Activity. *http://dx.doi.org/10.1080/10508406.2014.896254* 23, 4 (10 2014), 490–536. https://doi.org/10.1080/10508406.2014.896254

[34] Manu Kapur and Katerine Bielaczyc. 2011. Designing for Productive Failure. *http://dx.doi.org/10.1080/10508406.2011.591717* 21, 1 (1 2011), 45–83. https://doi.org/10.1080/10508406.2011.591717

[35] Royce Kimmons and George Veletsianos. 2018. Public Internet Data Mining Methods in Instructional Design, Educational Technology, and Online Learning

Research. *TechTrends 2018 62:5* 62, 5 (6 2018), 492–500. https://doi.org/10.1007/S11528-018-0307-4
[36] Simon Knight, Roberto Martinez-Maldonado, Andrew Gibson, and Simon Buckingham Shum. 2017. Towards mining sequences and dispersion of rhetorical moves in student written texts. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*. 228–232.
[37] Simon Knight, Alyssa F Wise, Bodong Chen, and Britte Haugan Cheng. 2015. It's about time: 4th international workshop on temporal analyses of learning data. In *Proceedings of the Fifth International Conference on Learning Analytics and Knowledge*. 388–389.
[38] Douglas Larkin. 2012. Misconceptions about "misconceptions": Preservice secondary science teachers' views on the value and role of student ideas. *Science Education* 96, 5 (9 2012), 927–959. https://doi.org/10.1002/SCE.21022
[39] Bruno Latour. 1986. Visualization and cognition. *Knowledge and society* 6, 6 (1986), 1–40.
[40] Amanda Lenhart, Kristen Purcell, Aaron Smith, and Kathryn Zickuhr. 2010. Social Media & Mobile Internet Use among Teens and Young Adults. Millennials. *Pew Internet & American Life Project* (2 2010).
[41] Nitin Madnani, Michael Heilman, Joel Tetreault, and Martin Chodorow. 2012. Identifying high-level organizational elements in argumentative discourse. In *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 20–28.
[42] Eve Manz. 2015. Representing Student Argumentation as Functionally Emergent From Scientific Activity:. *http://dx.doi.org/10.3102/0034654314558490* 85, 4 (12 2015), 553–590. https://doi.org/10.3102/0034654314558490
[43] Roberto Martinez-Maldonado, Yannis Dimitriadis, Alejandra Martinez-Monés, Judy Kay, and Kalina Yacef. 2013. Capturing and analyzing verbal and physical collaborative learning interactions at an enriched interactive tabletop. *International Journal of Computer-Supported Collaborative Learning* 8, 4 (2013), 455–485.
[44] Katherine L. McNeill, David J. Lizotte, Joseph Krajcik, and Ronald W. Marx. 2009. Supporting Students' Construction of Scientific Explanations by Fading Scaffolds in Instructional Materials. *http://dx.doi.org/10.1207/s15327809jls1502_1* 15, 2 (2009), 153–191. https://doi.org/10.1207/S15327809JLS1502{_}1
[45] Humphrey Mensah, Lu Xiao, and Sucheta Soundarajan. 2020. Characterizing the Evolution of Communities on Reddit. *ACM International Conference Proceeding Series* (7 2020), 58–64. https://doi.org/10.1145/3400806.3400814
[46] Matthew L. Newman, James W. Pennebaker, Diane S. Berry, and Jane M. Richards. 2016. Lying Words: Predicting Deception from Linguistic Styles:. *http://dx.doi.org/10.1177/0146167203029005010* 29, 5 (7 2016), 665–675. https://doi.org/10.1177/0146167203029005010
[47] NGSS Lead States. 2013. *Next Generation Science Standards: For States, By States*. Technical Report. The National Academies Press, Washington, DC.
[48] Ha Nguyen, Kyu Yon Lim, Liang Li Wu, Christian Fischer, and Mark Warschauer. 2021. "We're looking good": Social exchange and regulation temporality in collaborative design. *Learning and Instruction* 74 (2021), 101443.
[49] Alandeom W. Oliveira, Valarie L. Akerson, and Martha Oldfield. 2012. Environmental argumentation as sociocultural activity. *Journal of Research in Science Teaching* 49, 7 (9 2012), 869–897. https://doi.org/10.1002/TEA.21020
[50] Jonathan Osborne. 2010. Arguing to learn in science: The role of collaborative, critical discourse. *Science* 328, 5977 (4 2010), 463–466. https://doi.org/10.1126/SCIENCE.1183944
[51] Jonathan Osborne and Sue Collins. 2010. Pupils' views of the role and value of the science curriculum: A focus-group study. *http://dx.doi.org/10.1080/09500690010006518* 23, 5 (2010), 441–467. https://doi.org/10.1080/09500690010006518
[52] James W Pennebaker, Ryan L Boyd, Kayla Jordan, and Kate Blackburn. 2015. *The development and psychometric properties of LIWC2015*. Technical Report.
[53] James W Pennebaker and Anna Graybeal. 2001. Patterns of natural language use: Disclosure, personality, and social integration. *Current Directions in Psychological Science* 10, 3 (2001), 90–93.
[54] James W Pennebaker and Laura A King. 1999. Linguistic styles: language use as an individual difference. *Journal of personality and social psychology* 77, 6 (1999), 1296.
[55] Dilhan Perera, Judy Kay, Irena Koprinska, Kalina Yacef, and Osmar R Zaïane. 2008. Clustering and sequential pattern mining of online collaborative learning data. *IEEE Transactions on Knowledge and Data Engineering* 21, 6 (2008), 759–772.
[56] Richard E. Petty and Duane T. Wegener. 2016. Matching Versus Mismatching Attitude Functions: Implications for Scrutiny of Persuasive Messages:. *http://dx.doi.org/10.1177/0146167298243001* 24, 3 (7 2016), 227–240. https://doi.org/10.1177/0146167298243001
[57] Anna McLean Phillips, Jessica Watkins, and David Hammer. 2017. Problematizing as a scientific endeavor. *Physical Review Physics Education Research* 13, 2 (8 2017), 020107. https://doi.org/10.1103/PhysRevPhysEducRes.13.020107
[58] J. Piaget. 1972. Intellectual Evolution from Adolescence to Adulthood. *Human Development* 15, 1 (1972), 1–12. https://doi.org/10.1159/000271225
[59] PISA. 2017. *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic, Financial Literacy and Collaborative Problem Solving*. Technical Report. OECD Publishing, Paris. https://doi.org/10.1787/9789264281820-EN
[60] R Core Team. 2021. R: A Language and Environment for Statistical Computin. https://www.r-project.org/
[61] Günter Rohdenburg. 1996. Cognitive complexity and increased grammatical explicitness in English. (1996).
[62] Carolyn Rosé, Yi Chia Wang, Yue Cui, Jaime Arguello, Karsten Stegmann, Armin Weinberger, and Frank Fischer. 2008. Analyzing collaborative learning processes automatically: Exploiting the advances of computational linguistics in computer-supported collaborative learning. *International Journal of Computer-Supported Collaborative Learning* 3, 3 (9 2008), 237–271. https://doi.org/10.1007/s11412-007-9034-0
[63] Tim Rowland. 1995. Hedges in mathematics talk: Linguistic pointers to uncertainty. *Educational Studies in Mathematics* 29, 4 (1995), 327–353.
[64] Victor Sampson and Douglas Clark. 2009. The impact of collaboration on the outcomes of scientific argumentation. *Science Education* 93, 3 (5 2009), 448–484. https://doi.org/10.1002/SCE.20306
[65] Michael Scholz. 2016. R Package clickstream: Analyzing clickstream data with Markov chains. *Journal of Statistical Software* 74, 1 (2016), 1–17.
[66] Simon Buckingham Shum and Rebecca Ferguson. 2012. Social learning analytics. *Journal of educational technology & society* 15, 3 (2012), 3–26.
[67] Edward L. Smith, Theron D. Blakeslee, and Charles W. Anderson. 1993. Teaching strategies associated with conceptual change learning in Science. *Journal of Research in Science Teaching* 30, 2 (2 1993), 111–126. https://doi.org/10.1002/TEA.3660300202
[68] Kumar Bhargav Srinivasan, Cristian Danescu-Niculescu-Mizil, Lillian Lee, and Chenhao Tan. 2019. Content removal as a moderation strategy: Compliance and other outcomes in the changemyview community. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (11 2019), 163. https://doi.org/10.1145/3359265
[69] Iris Tabak and Eric Baumgartner. 2010. The Teacher as Partner: Exploring Participant Structures, Symmetry, and Identity Work in Scaffolding. *http://dx.doi.org/10.1207/s1532690Xci2204_2* 22, 4 (2010), 393–429. https://doi.org/10.1207/S1532690XCI2204{_}2
[70] Chenhao Tan, Vlad Niculae, Cristian Danescu-Niculescu-Mizil, and Lillian Lee. 2016. Winning arguments: Interaction dynamics and persuasion strategies in good-faith online discussions. *25th International World Wide Web Conference, WWW 2016* (2016), 613–624. https://doi.org/10.1145/2872427.2883081
[71] John C Turner, Michael A Hogg, Penelope J Oakes, Stephen D Reicher, and Margaret S Wetherell. 1987. *Rediscovering the social group: A self-categorization theory.* Basil Blackwell.
[72] Jessica Watkins, David Hammer, Jennifer Radoff, Lama Z. Jaber, and Anna M. Phillips. 2018. Positioning as not-understanding: The value of showing uncertainty for engaging in science. *Journal of Research in Science Teaching* 55, 4 (4 2018), 573–599. https://doi.org/10.1002/TEA.21431
[73] Zhongyu Wei, Yang Liu, and Yi Li. 2016. Is this post persuasive? Ranking argumentative comments in online forum. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. 195–200.
[74] Armin Weinberger and Frank Fischer. 2006. A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers and Education* 46, 1 (1 2006), 71–95. https://doi.org/10.1016/j.compedu.2005.04.003