Signals Beyond Text: Understanding How Accessing Peer Concept Mapping and Commenting Augments Reflective Mind for High-Stake Videos

Jingxian Liao
Department of Computer Science
UC Davis
Davis, California, USA
jxliao2021@gmail.com

Fu-Yin Cherng
Department of Computer Science and Information Engineering
National Chung Cheng University
Chiayi, Taiwan
fuyincherng@gmail.com

Mrinalini Singh
University of California, Davis
Davis, California, USA
mrisingh@ucdavis.edu

Hao-Chuan Wang
UC Davis
Davis, California, USA
hciwang@ucdavis.edu

Abstract

In high-stakes domains, deep analytical processing of online videos is essential for decision-making and knowledge acquisition. However, individuals may lack sufficient cognitive resources and triggers to engage in such processes. To address this, we introduce DeepThinkingMap, a collaborative video mapping system with affordances designed to leverage peers' thoughts and comments to promote reflective and critical thinking. Thee design supports collaborative mapping of video concepts and supports open deliberations of personal thoughts over concepts as "thinking nudges" to foster deeper thinking for themselves and others. Through two experimental studies, we investigated the potential of deeper thinking by accessing peers' thoughts in standalone and collaborative information work respectively. Results illustrated that accessing peers' comments enhances personal engagement in reflective and critical thinking, and reinforces their confidence in their correct beliefs toward the video topics. This work contributes to understanding the socio-technical-cognitive mechanism of thinking while accessing peer comments, and presents design implications for information and knowledge work.

CCS Concepts

• **Human-centered computing → Collaborative and social computing**; **Empirical studies in collaborative and social computing**; Visualization application domains.

Keywords

Thinking Nudge, Reflection, Critical Thinking, Higher-order Thinking, Concept Mapping

CHI '25, Yokohama, Japan

ACM ISBN 979-8-4007-1394-1/25/04
https://doi.org/10.1145/3706598.3713426

ACM Reference Format:
Jingxian Liao, Fu-Yin Cherng, Mrinalini Singh, and Hao-Chuan Wang. 2025. Signals Beyond Text: Understanding How Accessing Peer Concept Mapping and Commenting Augments Reflective Mind for High-Stake Videos. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26–May 01, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 21 pages. https://doi.org/10.1145/3706598.3713426

1 Introduction

In the era of online infodemics filled with unverified, unstructured content created by online crowds and generative AI, online content's credibility, quality, readability, and comprehensiveness of online content have become focal points of public concerns, especially in high-stakes domains, including healthcare and financial investments [39, 72]. Moreover, in addition to quality issues, general information available online is not personalized to individual needs. Yet, information seekers (e.g., those seeking diet suggestions for personal health conditions) often overlook such a mismatch of personal and general contexts, which can pose additional risks.

With the rapidly changing informational sphere surrounding information seekers and consumers [58], it is critical for individuals to engage in deeper deliberation and scrutinization of information. However, personal factors like cognitive bias, limited time and resources, and sociopolitical influences from opinion leaders often lead people to take mental shortcuts, not triggering the personal practice of engaging in the comprehensive analysis of online content [2]. Therefore, empowering people to transcend inertial thinking and consciously use their reflective minds to judge high-stakes content is essential for individuals' and communities' well-being and safety [43].

To process information beyond literal comprehension, individuals need to make substantial cognitive efforts to integrate, analyze, and reflect upon such information. Doing so involves extensive engagement in higher-order thinking or practices beyond memorizing information or retelling memorized stories [36]. Higher-order thinking encompasses reflective and critical thinking. Both are essential when deciding how the content is applied to everyday work and lives (e.g., whether to get vaccinated) [44]. Reflective thinking involves a series of cognitive operations that identify the underlying structure of external information and deliberately connect them from the external source to the individual's existing mental models [44]. In contrast, critical thinking, in the context of information work, typically refers to the mode of thinking attempting to seek and apply reliable reasons to scrutinize arriving information, allowing individuals to remain open to diverse perspectives, especially when encountering information that challenges their existing beliefs [62].

Despite the critical value of these cognitive endeavors in information work, higher-order thinking is often cognitively taxing and demands skill. Thus, levels of higher-order thinking vary across contexts and among diverse types of thinkers [38]. For example, people with a low score on the Need for Cognition (NFC) index—the measure of individual differences in engaging in thinking—are more susceptible to the influence of superficial messages or simple cues [22]. Such observation creates a tangible need to help people trigger their reflective minds. Inspired by the boosting effects of nudging strategies as behavior interventions [28] and displaying exemplary behaviors in augmenting cognition—such as when deliberating on political views or verifying the credibility of informational sources [50]—we note the potential to *engage individuals in deep thinking* through carefully designed external interventions. In the context of behavioral change, nudging refers to choice-preserving interventions that guide behaviors toward desirable states without constraining behavioral options [95]. When applied to cognitive activities, nudging can be considered as a form of less structured and less demanding socio-cognitive support than formally organized thinking protocols (e.g., reflective journaling) and may reduce the likelihood of burnout and resistance reactions [95, 98]. The open-ended and flexible nature of ***thinking nudges***, or nudging designs that aim to promote higher-order thinking, offers a promising approach to scaffold individual- and peer-based information work, which encourages the expansion of thinking efforts by keeping the space of thinking rich yet open.

Several studies have leveraged individuals' and groups' collective efforts and wisdom to facilitate information work of high-stakes content, such as verifying its authenticity, curating group discussion and design annotations, and reorganizing information from different sources [5, 24, 31, 51, 65, 105]. Existing system support for collaborative information work varies in several key aspects, such as the source of support (e.g., from experts, strong-tied small groups, online community members, or AI models) [12, 69, 70], the mechanism employed to provide support (e.g., providing straightforward recommendations or raising users' awareness) [65, 68], organizing how resources are aggregated (e.g., only showing selected items or visualizing most resources available) [16, 107], and the timing of the support provided [24]. While many of these designs aim to provide cognitive tools conducive to thinking, few have explicitly focused on fostering higher-order thinking when engaging with high-stakes content on generic platforms that may have mixed user-generated and authoritative content.

In this work, we take an alternative route to support high-stakes information consumption. We hypothesize that observing peers' conceptualizations and commenting on the same information in generic platforms may provide social behavioral cues ("Do others reflect as well?") and cognitive stimulation ("How are others responding to this content?"), thereby motivating and fostering engagement in higher-order thinking. Interactions between peer users in these scenarios are usually informal and loosely connected, sometimes asynchronous and fragmented [90]. Therefore, to highlight the behavioral cues and cognitive stimuli for thinking nudge, one design challenge lies in explicitly representing and displaying peers' thinking processes [24, 59]. Reflective and critical thinking are cognitively demanding operations that individuals may struggle to engage with and externalize. The complexity in the design arises in collaborative information work, which requires team members to share their thoughts while considering others' ideas, especially when group members have no grounded understanding of specific information. This situation requires anchoring others' thoughts on the original informational content and mitigating the possible dominance of popular or recent opinions that can lead to groupthink and hinder deep thinking [56].

To address these challenges, we designed a collaborative video mapping interface, DeepThinkingMap, that supports visually segmenting and mapping video content and sharing personal insights to aid the digestion of high-stakes materials, such as videos in the healthcare domain (see Figure 1). The interface invites individuals to collaboratively identify the embedded concepts of the video content and connect these concepts to form semantic structures [26, 54, 99]. Additionally, the interface features a peer commenting feature that prompts peers to share their thoughts beyond mere concepts or content summaries. These personal comments are anchored to the original material and can be linked to other comments and conceptual highlights in a non-linear and multi-directional fashion, helping preserve the diversity and complexity of thoughts and encourage individuals to make non-trivial, insightful connections between diverse thoughts [101].

We conducted two lab-based studies to investigate the viability of the interventions in solitary and collaborative information work. The studies are intended to validate the system mechanism before field deployment. Study 1 focused on the impact of accessing peer comments on personal engagement in thinking in stationary and solitary work. Study 2 further studied collaborative information work using the video mapping interface, DeepThinkingMap, particularly in considering the effects of concept mapping organization and presentation of peer comments. In Study 1 ($N = 44$), notes of median quality from a previous user were used as cues to investigate the general effects of thinking nudges in a controlled setting where the cues are prescribed and static. The results revealed that merely viewing average-quality notes from a prior user can stimulate reflective thinking, demonstrating the potential of asynchronous comment sharing for deeper thinking. Study 2 ($N = 52$) expanded on the approach used in Study 1 by evaluating not only the influence of thinking nudges in a synchronous interactive setting but also by comparing DeepThinkingMap with a conventional linear-fashion word-processing tool, Google Docs. Google Docs was chosen for comparison as one widely used tool for collaborative information work. This comparison allowed us to see how thinking-nudge affected users with and without interactive features designed for the process. The results from Study 2 corroborated the findings of Study 1, highlighted the significant impact of DeepThinkingMap on critical thinking and attitudinal changes. The consistency of the studies' results indicates that accessing peers' understanding and personal comments can (1) enhance but not impede engagement in reflective and critical thinking, (2) elevate the analytical depth of thoughts via peer nudging, and (3) bolster users' confidence in their attitudes to a topic, which can differ from online information. Furthermore, the impact of thinking nudges was amplified by the concept-mapping type of documentation and visualization spaces afforded by DeepThinkingMap. Our main contributions are as follows:

- Demonstrating the socio-technical-cognitive mechanisms of how accessing peer concept mapping and commenting can subtly encourage reflective and critical thinking in video reviewing.
- Generating empirical understandings and new insights from lab studies on thinking-nudge designs, including the benefits of non-linear visualizations, and how risks of negative social processes like social conformity can be mitigated.

2 Background and Related Work

2.1 Higher-order Thinking

Higher-order thinking is crucial in the era of information proliferation, as it allows continued creativity and solutions [53]. According to Lewis and Smith, higher-order thinking occurs when a person takes new information and information stored in their memory and integrates and extends them to achieve a purpose [66]. Online information consumption often requires higher-order thinking, including decision-making for content selection, integrating diverse sources, creating a coherent mental model of the topic, and evaluating opposing viewpoints [44]. We adopt Geertsen's framework to explore how thinking nudges may facilitate two main types of higher-order thinking: critical and reflective thinking [44]. Reflective thinking involves activities that connect different thoughts and enlarge the space for thinking, such as analyzing the logical relations between ideas for sense-making and finding possible similarities between the known and the newly learned. Critical thinking consists of establishing the accuracy and validity of content through self-examination, such as critically evaluating a news article's sources to determine its credibility. Despite the distinctions in their definitions, critical and reflective thinking are both vital for processing high-stakes online information that may include unverified content [43]. For example, for diet suggestions, people must assess alleged facts that support the benefits of certain diet strategies (i.e., critical thinking) and compare them with their existing health knowledge (i.e., reflective thinking).

The HCI community has shown significant interest in supporting higher-order thinking and the meta-cognitive process of managing one's own thinking process [19], particularly reflection [10, 11]. Previous studies have focused on understanding and supporting reflective practices in different contexts [10]. Examples include reflecting on health habits using personal activity data (personal informatics) [51], re-processing knowledge for new conceptual schemas (education) [78], and encouraging reflection in group creativity work (creativity) [30]. There have also been attempts to identify design principles for supporting reflections [14]. Tools like NewsCube [82] and Trackly [7] demonstrate how clustering viewpoints or using customized informational trackers can prompt reflective thinking. Social discomfort has also been explored as a means of promoting critical thinking [52], and TalkReflection supports collaborative reflection through experience sharing [85].

To summarize, prior work has focused on enhancing reflective practices in specific scenarios and contexts. The question of supporting higher-order thinking while preserving individual choice and incorporating access to peer thought processes remains underexplored. Given the prevalence of collaborative information work, the emphasis is shifting to understanding the cognitive impact of openly sharing and comparing individual and peer thought processes.

2.2 Thinking-Nudging in Knowledge Work

With the emergence of personal informatics and collaborative computing, the HCI community has been interested in what design interventions can offer choice-preserving options for desirable changes in personal and social contexts, such as health informatics, recommendation systems, and news consumption [24]. While there are nuanced differences in how such interventions should be conceptualized and operationalized, *nudging* is generally seen as a lightweight intervention directing people toward particular choices while preserving all options [93]. The initial conception of nudging was conceived to occur without much conscious awareness and would require only low cognitive capacity for people to shift their *behaviors* [95]. As the framework and questions around behavioral change developed over the past decade, *research on nudging advocates the use of both cognitive and reflective elements to improve both behavioral and thinking outcomes* [8]. There is also an ongoing discussion concerning the efficiency of nudging on behavior policy [28, 73], which provides meta-review results from the broad grouping of diverse goals, populations, and contexts under the umbrella of nudging. However, not all designs need or aim to re-formulate people's behavioral policies. This work focuses specifically on using peers' sharing and commenting behaviors as a *thinking nudge*, which provides choice-preserving opportunities for knowledge workers to process the information deliberately and reflectively.

Researchers have increasingly investigated how nudge-inspired design interventions can encourage and guide sense-making and decision-making in online environments [15, 24]. Caraban et al. [24] summarize 23 nudging mechanisms in six categories: facilitate, confront, deceive, fear, reinforce, and social influence. These mechanisms echo the design approaches for reflection support discussed above. Our work echoes the mechanisms related to social influence. For example, personalized messages and peer comparisons have been shown to encourage students to reflect on study habits and complete assignments [21]. Gimpel et al. found that exposing users to a mix of truthful and misleading social media articles increased reflective comparison [47]. For active video exploration, designers used social tagging and comment ratings to provide crowd-referenced resources to nudge more deliberate video selection [31, 65]. These signals usually provide socially grounded references to how others perceive and behave, typically normalizing individual choices.

While nudging to think and change can be helpful, its effectiveness often depends on the context of specific user groups. For example, there have been mixed study results when users receive nudging interventions from groups that differ significantly from their own [32, 71]. When a reference group is irrelevant to the target individuals, disseminating information about that group's behavior could have a limited effect [17]. Moreover, social psychological studies on group work have implications for the potential pitfalls. Individuals' desire for social conformity with mainstream ideas can promote alignment with peers, and fear of judgment may lead to decreased diversity of thoughts, resulting in groupthink. Putting individuals in team settings to perform work that is not individually accountable can also lead to social loafing or excessively relying on other team members to do the work [29, 56]. However, it remains unclear whether accessing peer comments in solitary or collaborative work for thinking activities would suffer from the known social pathology, especially as past studies on nudging interventions give clues to the potential benefits of nudging for thinking. The current work aims to contribute to this design space with our nudging interface and research examining how peer comments could prompt thinking.

2.3 Socio-cognitive Resources for Learning and Thinking

2.3.1 Social Annotation for Learning. Social annotation platforms are gaining traction for supporting higher-order thinking in online learning through collaborative highlighting and discussion of learning materials. These platforms have been studied from multiple perspectives, including system design for content curation and visualization [5, 49] and their impact on reflection and comprehension [18, 102]. This section discusses common social annotation designs and connects them with DeepThinkingMap.

Social annotation tools provide methods for sharing, visualizing, and interacting with learning content to enhance discussion and collaborative learning. Users can highlight text and add, respond to, or upvote comments [5, 37, 49]. Platforms also use navigation aids like sidebars, hyperlinks, and visual techniques such as color-coding and semantic grouping to improve peer interaction and focus attention [67, 69, 92]. With the rise of video-based learning, features such as video transcript annotation and progress bar visualization have emerged [37, 92]. DeepThinkingMap incorporates visualization techniques from social annotation and video-based discussions, using a canvas positioned alongside video content to differentiate notes with color and link timestamps to peer contributions, enabling seamless navigation. While social annotation tools typically center on expert-sourced materials within specialized educational platforms, this work broadens this scope by supporting individual and collaborative engagement with high-stakes video content on generic content-sharing sites. These video platforms (e.g., YouTube) are filled with both user-generated and expert-generated content. Moreover, many educational YouTube videos include entertaining elements to captivate audiences, blurring the lines between traditional education and entertainment. Thus, in this work, we move beyond the conventional boundaries of social annotation systems to address more diverse and open-ended informational tasks.

2.3.2 Concept Mapping and Concept Map Visualization. Concept mapping represents the structure of domain knowledge as an external graph of concepts and their interconnections for a specific topic (e.g., vaccination) [79]. Concept mapping can be effectively applied to abstract topics and concrete content related to those topics (e.g., a video introducing the topic of vaccination). Concept mapping has been broadly applied in education and can serve various learning-related purposes. As a learning exercise, the technique encourages deep thinking by requiring students to transform complex and nuanced conceptual information into concrete visual representations. Concept mapping can also be used to assess students' levels of learning. Concept mapping is also a metacognitive activity that requires students to filter out irrelevant information, summarize critical concepts, and understand how concepts intersect in the domain [54]. Collaborative concept mapping encourages more group discussions and helps identify different perspectives [97], which may scaffold deeper reflection and critical analysis of the content.

For visualization, concept maps are an alternative representation that helps communicate information [23, 42]. Traditional sharing and collaboration tools present information in a linear format, such as threaded forums or Google Docs. They may cause recency and primacy effects that can skew attention [48]. People's tendencies to recall the most recent or popular content may hinder the impact of divergent thoughts on individuals across different interaction timelines. In contrast, concept maps offer flexibility in how ideas are linked and how the connections relate. The spatial and relational layout of concepts may serve as retrieval cues to help individuals shift the focus from the positions of content to its relevance and connection to other ideas. Moreover, concept maps may help spread attention across nodes and counterbalance the serial-position effect (i.e., the tendency to attend to or remember items at the beginning and end of a list better than items in the middle) [9]. Prior studies in HCI have applied concept mapping to support informal video learning [69], facilitate document processing and comprehension [106], and facilitate knowledge integration among complex information sources [100]. While representations and mechanics of concept mapping are utilized in this work, we focus on supporting simultaneous content mapping and thought-sharing as individuals and peers watch and discuss video materials.

3 Thinking-nudge and DeepThinkingMap Design

3.1 Mechanisms

Our intervention design is rooted in the notion of increasing the transparency and accessibility of higher-order thinking cues in terms of peers' commenting behaviors and the actual comments peers produced [55]. Engaging in higher-order thinking is cognitive, behavioral, and motivational, as individuals may choose not to make the effort unless they are appropriately triggered [22].

Research on social influence suggests that watching peers engage in cognitive tasks can serve as thinking nudges, encouraging individuals to make similar efforts voluntarily [27, 45]. This tendency is independent of the quality of peers' thinking outputs [25]. Such sharing may tap into people's social motivators by providing opportunities for social rewards, such as acknowledgment or a sense of community. For instance, when there is visible engagement in thinking activities in the community, people are more likely to engage in similar efforts to follow the social norm [89]. To harness these social mechanisms, the interface is designed to make peers' thinking activities transparent by making how others process the online information vis-à-vis the original material visible. Building on these act cues, the content of peer comments may have a secondary impact on thinking. Peers' thought-sharing increases the likelihood that individuals will encounter information that challenges their existing mental models. It can lead individuals to reassess the videos and engage in higher-order thinking activities to resolve the conflicts [96]. Multiple pathways exist through which peer commenting behaviors and content can influence thinking processes.

Regarding the sources of intervention, we wonder how pre-generated peer comments displayed in solitary work and interactive peer comments emerge through peer collaborations that impact deep thinking. Specifically, in Study 1, participants worked alone. We collected peer comments offline and selected the comments from participants whose products were rated as medium coverage of video concepts as thinking nudges. In Study 2, participants collaborated in three-person teams and shared their comments while accessing others' comments in real-time. As comments are essentially from the same group, the reference group is identical to themselves, which ensures the transferability and applicability of thinking nudges [17].

Another design consideration is how to mediate peer-sharing cues. Videos have become a key medium for content sharing and dissemination [75, 86]. However, user-generated videos are not always ideal for reflective and critical thinking. For instance, comparing information across different sections of the same or different videos is challenging without additional work and support for handling timestamps. Moreover, presenting audience reactions and comments generated by various individuals across video timestamps is difficult. A traditional linear design that displays user comments as a chronologically ordered list can hinder the visibility of older comments [90] and further reduce their utility in nudging for thinking.

3.2 Design and Implementation of DeepThinkingMap

3.2.1 Design Principles in DeepThinkingMap. DeepThinkingMap is designed as a non-linear concept-mapping tool for videos that supports creating a graphical network to capture the conceptual highlights embedded in a video. DeepThinkingMap differs from conventional concept mapping by allowing users to express personal insights and anchor conceptual summaries to specific video timestamps where the concepts are introduced. Users can also add comments beyond the original materials and connect them to the relevant concepts and video sections in the canvas. The spatial arrangement of information on the canvas allows users to interact with content in a non-linear manner, mitigating biases such as status quo and anchoring, which are often reinforced by traditional, sequential text-based formats like word-processing documents and threaded comments [15, 24, 63, 97]. By externalizing thoughts onto a canvas, concept mapping can enhance the clarity and organization of internal ideas. These activities can offer substantial benefits in grounding people's comprehension of the original content and encouraging higher-order thinking with the behavioral cues and cognitive stimuli they supply.

Based on the considerations discussed above, we outline three design principles (DP) for DeepThinkingMap:

DP1: Fostering thinking transparency. To trigger higher-order thinking, it is crucial to make the presence of peers and their cognitive activities accessible to users [14]. The interface should explicitly highlight the social presence of others and their cognitive actions, and also allow users to compare their thoughts with those of their peers efficiently.

DP2: Encouraging and differentiating video content and personal thoughts. Higher-order thinking extends beyond the original content of videos. The interface should clearly differentiate between the two to ensure users' cognitive contributions are distinct from mere extraction or rephrasing of video content. Furthermore, the interface should provide clear prompts encouraging users to generate personal insights.

DP3: Presenting temporal connections between personal notes and video content. Since users often anchor their comments to specific video segments, the interface must seamlessly integrate these notes with the corresponding content. This design should allow two-way referencing—links between notes to particular video segments—facilitating a deeper understanding and interaction with the content and peer contributions [103].

3.2.2 Example Usage Scenario. Mary is a health enthusiast who wants to watch YouTube health videos to better understand the latest wellness trends and peer audiences' reactions. After opening DeepThinkingMap, she noticed that it provides a video player and a large, interactive canvas on the same screen (Figure 1). She sees that the video player ① has essential controls—play, pause, and the ability to jump around in the video—so she can watch at her own pace. She also realizes she can navigate among multiple pre-selected videos by using dedicated tabs ②, quickly switching from one health topic to another.

Next to the video player, Mary finds the collaborative mapping canvas ③ where she can jot down her insights and read what others have shared in real time. Two input boxes await her contributions ④: one labeled "Add/edit video excerpt," which is used for summarizing specific segments in the video, and one labeled "Add your post," for sharing her own thoughts and ideas (**DP2**). A button labeled "Hide comments" ⑤ allows her to toggle whether to see only these short video summaries or to also view personal posts from herself and others. This way, she may choose to form her own understanding of the material before looking at others' perspectives.

When Mary starts taking notes, she notices that the canvas uses shapes and colors to help her distinguish different kinds of contributions. Video excerpts appear as light-yellow rectangles, while personal comments appear as violet ellipses ⑦. These visual elements are aligned with the corresponding input box colors (**DP2**). Mary also has the option to draw directed links between notes, which is helpful for connecting two related excerpts or clarifying the relationship of her personal comments to specific parts of the video (**DP1**). Whenever Mary hovers over a note ⑥, it becomes highlighted, and the system shows the username of the last person who edited it, providing a subtle but reassuring sense of collaboration.

Because the system automatically logs the video's current timestamp whenever Mary creates a note, she can simply double-click on any note to jump straight to the moment in the video it refers

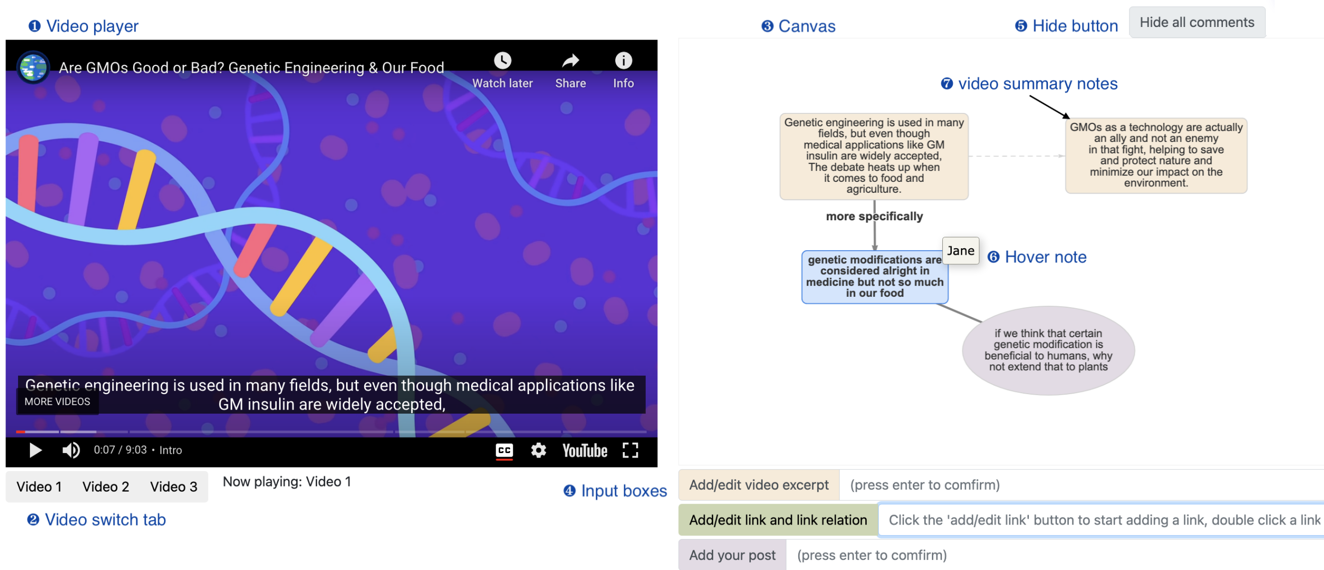


Figure 1: DeepThinkingMap Interface. The tool includes a standard video player with a video switch tab on the left, allowing for in-tab video switching. The note canvas for real-time note-taking and discussion is on the right. Input boxes for video summaries and individual comments are below the canvas. Additional features like the "hide comments" button and hover-over identity indicators facilitate collaboration.

to (**DP3**). It helps her track which notes go with which parts of the video, making her video digest more structured and immersive. When Mary finishes watching a few clips, her canvas is filled with color-coded summaries and personal thoughts, neatly organized with helpful links. She leaves the session feeling better informed about her health topics of interest and engaged by DeepThinkingMap's easy sharing and collaboration.

3.2.3 Implementation Notes. The DeepThinkingMap prototype is implemented as a web application for the discussions around the videos. It utilizes the YouTube Data API for client-side video sourcing and D3.js for interactive functionalities. It uses the Firebase database for real-time synchronization [1].

4 Study 1: Nudging Thinking in Non-Interactive Individual Work

In Study 1, we investigate how the visibility of peer thinking activities influences personal engagement with higher-order thinking activities of videos with DeepThinkingMap in standalone information work.

4.1 Study Design and Hypotheses

We designed a standalone information work scenario where we could manipulate the visibility of peer comments and control the content of these comments for all participants. For this purpose, we invited two batches of participants to finish a video-watching task in sequence. The first batch of participants watched a video with a blank canvas and wrote down their notes (control condition); the second batch of participants finished the same task with notes of one prior participant selected from the first batch as thinking nudges (nudge condition). It is a between-subject study comparing the influences of others' thoughts on engagement in thinking activities and notes.

The focus of the study is on verifying the core hypothesis of thinking nudges—whether seeing peer-generated thinking can prompt oneself to exercise higher-order thinking. We hypothesized that accessing peer thoughts would improve video comprehension and engagement in higher-order thinking activities (i.e., critical and reflective thinking). We tested this hypothesis by comparing two conditions—*with* versus *without* access to peer-generated comments when watching the video and thinking about the video content.

4.2 Participants

We recruited 44 participants (28 female, 14 male, and 2 non-binary), aged 19-29, from one university campus after pre-screening for common healthcare literacy, independent ability to make health decisions, and English fluency. Two groups of participants were enrolled: 24 participants for the control group and 20 in the nudge condition. They self-reported healthcare knowledge as median level ($mean$ = 4.93, sd = 0.99, on a 7-point scale where 1 = "novice level" and 7 = "expert level") based on the questions from [57] (e.g., "I am familiar with preventing minor and chronic problems such as allergies and dry skin."). As background knowledge is pivotal to reflective and critical thinking [35], we compared healthcare knowledge scores between participants from different conditions and found no significant difference ($t = -0.15, p = 0.87$). The study lasted about 30 minutes, and participants received about $7 USD as compensation. The study was approved by the institution's ethical review board.

4.3 Material Preparation

4.3.1 Video Materials. We selected two healthcare videos from YouTube. One introduces the human immune system ("Immune System"), explaining multiple lines of immune defense against bacteria. The other talks about the scientific evidence about the influence of turmeric in golden lattes ("Golden Latte"). These videos come from high-profile YouTubers whose popularity suggests broad visibility, with no indicator of health domain expertise. Appendix A shares the metadata of these videos. We chose these introductory videos to reduce the burden of specialized knowledge on participants' comprehension. We shortened the videos to approximately five minutes, to ensure enough time for thinking and writing notes. About 95% of the participants were unfamiliar with their assigned video, while some already had background knowledge. Participants were randomly assigned to one of the two videos.

4.3.2 Content Construction. To prepare the content for the nudge condition, we collected notes from participants in the first batch. These 24 participants were invited to perform the task (i.e., the control condition). Given a blank canvas on DeepThinkingMap, we instructed participants to create their personal notes, including conceptual takeaways from the video and personal comments (e.g., opinions, related experience, and questions) with different input boxes. We also encouraged participants to link different notes with links and labels, demonstrating the connections they saw.

There were 12 individual works completed by participants from batch 1 for each video, as the 24 participants were randomly assigned to one of two videos. To control the quality of thinking nudges across different videos, we chose two individual notes of average quality from each video as the content of thinking nudges. Two researchers independently calculated the number of unique concepts mentioned in each video, finding 28 concepts for the "Immune System" and 36 for the "Golden Latte." We then calculated the conceptual coverage (i.e., *#concepts_mentioned* / *#concepts_ever_mentioned_by_anyone*) for each participant's notes and chose those whose coverage ranked at the median. As we selected the content by the coverage of video content, we displayed the full notes these four participants wrote. In summary, we selected four participants from the first batch and used their personal comments to influence the second batch of participants. These four participants and their data were excluded from the subsequent analysis that compares the two conditions to avoid possible confound.

4.4 Procedure

We conducted Study 1 via our DeepThinkingMap website. Upon recruitment, participants received an invite link to complete tasks and questionnaires online. We first introduced the study procedure and gave them a tutorial on the task and DeepThinkingMap. Depending on the study condition, DeepThinkingMap's canvas either pre-displayed the thinking-nudge content before the work started or remained empty. Participants were asked to watch the video, jog down conceptual highlights to summarize the video, and add personal comments on the canvas of DeepThinkingMap. In the nudge condition, the instruction was adjusted to match the availability of thinking-nudge content. We shared with the participants that the existing notes were generated by a peer participant, and they are welcome to add their own sharing. We imposed a 20-minute limit for the task in both conditions. Upon task completion, participants completed a post-study survey on their engagement levels in different types of thinking activities.

4.5 Measures and Data Analysis

4.5.1 Engagement in Different Thinking Processes. In the post-task survey, participants rated their reflective and critical thinking engagement using a 5-point Likert scale (1 – strongly disengaged, 5 – strongly engaged). The survey items, adapted from Kember et al. [60], are provided in the appendix. We also asked about participants' engagement level of understanding as a comparison [61]. Cronbach's alpha is 91.6%.

There are other HCI-oriented measures assessing thinking processes in interactions, such as the Technology-Supported Reflection Inventory (TSRI). It is a scale measuring how effectively a system supports reflection from insight, exploration, and comparison [13]. Notably, TSRI shares similar conceptualizations and operationalizations with Kember et al.'s survey [60]. Both instruments identified different aspects of thinking, such as tackling challenges to existing understandings, reflecting on experiences, and considering alternative perspectives. TSRI primarily aims to evaluate the effectiveness of systems, while Kember's targets the underlying cognitive processes independent of systems. We opted for Kember's, which is in alignment with our research goal and framework.

4.5.2 Behavioral Logs. To understand how participants actively watched the video and wrote notes, we logged video player interactions and edits on the canvas. We tallied video pauses and rewinds as indicators of active video watching and quantified individual canvas operations—adding, editing, or deleting nodes and edges—to measure note-writing participation. These logs serve as proxies for participants' activeness in higher-order thinking.

4.5.3 Comment Category Analysis. Like reflective journals [84], participants' comments reveal the depth of higher-order thinking using code schemes as another metric to understand the effect of thinking nudges. There are many alternative measures to evaluate the thinking activeness directly, such as the think-aloud protocol. However, think-aloud may introduce more cognitive burden to participants [81], which requires training and may not be applicable in studies that demand naturalistic and social interactions [33]. As a trade-off, we analyze the notes participants left to understand their thinking depth.

Two coders first identified the personal comments by marking notes not explicitly mentioned in the video. Then, the coders classified these comments into three cognitive categories based on Anderson and Krathwohl's revision of Bloom's taxonomy [6]: understanding, analysis, and evaluation. The categories of analysis and evaluation represent the level of reflective thinking and critical thinking. The coders reached a high level of agreement, with a Cohen's Kappa > 0.9.

4.6 Results

To evaluate the influence of thinking nudges on engagement in different thinking processes, we conducted a mixed-effect one-way analysis of variance (ANOVA). The main independent variable is the study condition (control vs. nudge). The video topic and the specific

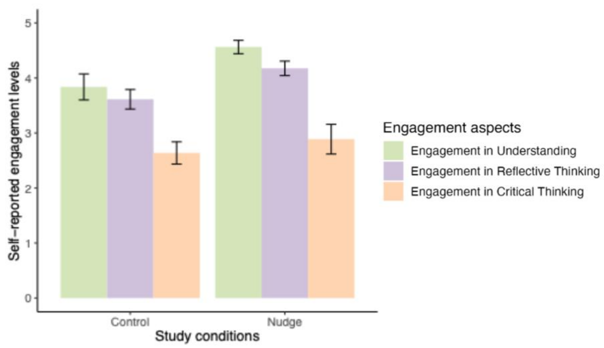


Figure 2: Level of Engagement in Thinking in Study 1. Error Bars Indicate the Standard Errors.

video-related thinking nudge notes were considered categorical random variables. Participants' self-reported score of healthcare knowledge was included as a confounding variable. For post-hoc analysis, we reported omega square (ω^2) for significant or near-significant effect sizes [104]. As the sample size of each condition is relatively small, omega square is a more unbiased effect size statistic compared to eta squared (η^2) [4, 104]. Statistical results are reported in Table 1.

4.6.1 Engagement in Thinking. For engagement in reflective thinking, ANOVA results show a significant difference between the control and nudge conditions in reflective thinking shown in Figure 2 and Table 1 (row 2). Moreover, ANOVA shows a trend toward significance in healthcare knowledge on reflective thinking engagement with marginal statistical significance.

For engagement in critical thinking, no significant differences were observed between the two conditions, as shown in Table 1 (row 3). Similarly, healthcare knowledge shows no main effect on critical thinking engagement.

To further investigate whether the difference between reflective and critical thinking is evident at the individual level, we performed a paired t-test to compare the two engagement scores for each participant. It shows that participants were significantly less engaged in critical thinking than in reflective thinking overall across both conditions ($t(39) = 3.76, p < 0.01$). Participants found it more difficult to engage in critical thinking, which is consistent with prior work [60].

As reported in Table 1 (row 1), the thinking nudges significantly increase engagement in understanding. Additionally, participants with a high score in healthcare knowledge were statistically more engaged in understanding the video ($\beta = 0.26$).

4.6.2 Commenting and Behavioral Logs. We first analyzed participants' behavior logs at the group level. To understand how their actions connect to the activation of higher-order thinking, we also summarized individual participants' behavior logs and associated the data with measures of thinking engagement.

Table 3 shows a significantly lower frequency of map editing behaviors in the nudge condition ($F(1, 37) = 25.13$, $p < 0.01$, $\omega^2 = 0.38$). The distributions of map editing in both conditions are skewed with a long right tail (Figure 3b). Participants influenced by peers' notes edited the map less frequently as some peer-generated conceptual summaries have been added to the canvas, reducing the need to duplicate these concepts. However, while the two conditions share similar video-viewing behaviors, participants in the nudge condition were able to engage in higher-order thinking and produce higher percentages of evaluative and analytical comments out of fewer map edits. There were a total of 37 comments generated in the control condition and 28 comments in the nudge condition. Figure 3a further shows the average number of comments generated by each participant in both conditions. While the raw counts of comments were lower in the nudge condition, over 50% comments in the nudge condition corresponded to results of reflective thinking and critical thinking (i.e., analysis and evaluation), compared to only 37% in the control condition.

When comparing behavior logs and participants' engagement in higher-order thinking at the individual level, we found statistically significant connections between map editing frequency and their commenting behavior in the control condition. With an empty canvas, the more actions participants performed on the canvas, the higher their engagement in understanding and reflective thinking, but no connections were found for critical thinking. Detailed statistics are presented in Table 2.

In summary, Study 1's results partially supported our hypothesis. Accessing peer comments as thinking nudges improved both the engagement in understanding and reflective thinking. However, there is no significant difference in the engagement in critical thinking in this individual work scenario. The analysis of behavioral logs provided more details about the connection between behaviors and engagement in thinking. With these results, we still need to understand whether the observed higher engagement in reflective thinking and proportionally better productivity in thinking outputs would hold in interactive settings when there are real social interactions and collaborations among peers.

5 Study 2: Nudging Thinking in Interactive Multiparty Work

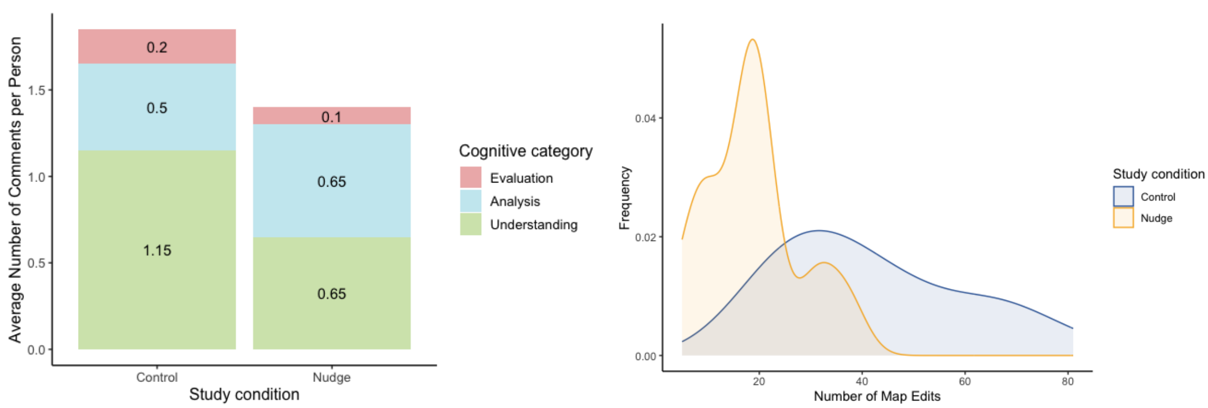
Study 1 explored the impact of DeepThinkingMap on higher-order thinking in video consumption. It provided insights into the role of thinking nudges on various thinking processes in an asynchronous setting for individual information work. Progressing from there, Study 2 aimed to deepen our investigation with dynamic and interactive thinking nudges in three-person collaborative work. During the collaboration, participants will witness peers creating video-related notes in real time and read the content synchronously. To understand the potential role of a non-linear mapping tool for thinking, we evaluated the impact of DeepThinkingMap by comparing it with conventional collaborative documentation tools (e.g., Google Docs). Note that collaborative documents like Google Docs are among the most widely and commonly used tools in everyday information work, thus serving as a reasonable baseline for comparison toward ecologically congruent understanding beyond using only the DeepThinkingMap tool developed in the lab. Through Study 2, we aim to provide a richer, more nuanced understanding of how people engage in higher-order thinking in collaborative information work with and without accessing peers' thinking activities.

Table 1: Summary of ANOVA Results for Thinking Engagement in Study 1.

Engagement in	Condition and confounding variable	F stats	p value	ω^2
Understanding	Control vs nudge	$F(1,37) = 7.75$	$p < 0.01^{**}$	0.15
	Healthcare knowledge	$F(1,37) = 3.94$	0.05	
Reflective Thinking	Control vs nudge	$F(1,37) = 6.59$	0.01^{*}	0.13
	Healthcare knowledge	$F(1,37) = 3.94$	0.05	
Critical Thinking	Control vs nudge	$F(1,37) = 0.26$	0.70	
	Healthcare knowledge	$F(1,37) = 2.27$	0.14	

Table 2: Correlations among Map Editing Frequencies and Engagement in Different Types of Thinking in Study 1.

Study condition	Measure 1	Measure 2	correlation coef	p-value
control	map editing	engagement in reflective thinking	0.51	0.02
control	map editing	engagement in understanding	0.55	0.01



(a) Number of Comments by Cognitive Categories.

(b) Distribution of Map Editing Frequencies by Conditions.

Figure 3: Comments and Map Editing Frequencies in Study 1.

Table 3: Means and Standard Deviations of the Frequencies of Video and Canvas Operations.

Condition	Map editing	Video rewind	Video pause	Comment score
Control	42.05 (18.39)	1.45 (1.47)	11.35 (10.83)	1.9 (1.62)
Nudge	18.75 (9.15)	1.85 (1.23)	17.85 (10.77)	1.35 (1.35)

5.1 Study Design

Study 2 is a between-subject experiment. The experiment is designed as a 2×2 factorial design: control (watch the videos individually) vs. nudge (in a 3-person group synchronously), and Google Docs (conventional collaborative tool) vs. DeepThinkingMap (graphical thinking nudge interface). Table 4 outlines the conditions and their abbreviations. Participants either completed the experiment individually or in triads using designated tools. In the nudge conditions, participants did not know each other beforehand. They also did not disclose their healthcare expertise or background to other participants.

5.2 Procedure

We used a similar task to Study 1, i.e., healthcare video consumption. We instructed participants to perform the video review and mapping task collaboratively. Since we noticed that watching single videos, as in Study 1, may not provide rich and complex enough information that would warrant multiparty collaboration, we increased the complexity of the task by inviting participants to review multiple videos sampled from different sources on the same topic, which also demands more higher-order thinking, especially in terms of critical thinking [44]. We set the word limit for each input up to 150 words (similar to the length of a tweet) in DeepThinkingMap to allow for more detailed contributions. This expansion allows participants to

Table 4: Study Conditions in Study 2.

	Control	Nudge
Google Docs	Individual video review using Google Docs (Control-Docs)	Video review using Google Docs in group (Nudge-Docs)
DeepThinkingMap	Individual video review using DeepThinkingMap (Control-DeepThinkingMap)	Video review using DeepThinkingMap in group (Nudge-DeepThinkingMap)

express their thoughts fully, fostering deeper discussions and more nuanced ideas.

We first introduced the task and played a pre-recorded tutorial to help participants familiarize themselves with the task and tools. Then, participants had 60 minutes to watch and review three videos about a randomly assigned topic. The topic assigned is either about the characteristics of Genetically Modified Organism foods [GMO] or how vaccination works [vaccination] (detailed descriptions of videos are in section 5.3). Specifically, participants need to review these three videos on the same topic holistically using their heuristics evaluation and work on their responses to an overarching question posed about the assigned video topic, "Do you think it is safe to vaccinate eligible people?" for the vaccination topic or "Do you think GMO food is safe?" for the GMO topic. As in Study 1, participants were encouraged to read and share their personal thoughts, review peer comments, and write down their notes on the videos. We surveyed participants twice about their personal attitudes and the confidence of their stances towards the overarching questions posed around the assigned topic: once before watching the video and once after the video review. They completed the same thinking engagement scale used in Study 1. The experiment ended with a semi-structured interview for about 5–15 minutes. Study 2 took about 90 minutes, and we compensated participants with a $15 gift card for their time.

5.3 Material Preparation

To situate the study in topics of potential controversy and vulnerability to misinformation, we used two healthcare topics frequently appearing in public discourse: GMO foods and general vaccinations. Despite broad scientific consensus on these two topics, online debates and discussions about their safety and impact remain, and widespread misunderstandings still threaten people's well-being. We curated 5–8 popular videos from YouTube for each topic and invited two professionals holding PhD degrees in pharmacology or public health communication to assess the veracity of the videos. Ultimately, we selected three videos per topic that align with the current scientific consensus. They are either from well-known YouTubers—whose substantial followings suggest broad appeal—or official university channels recognized for their academically rigorous content. By adopting the videos consistent with scientific agreement, we aimed to ensure participants' safety and avoid any ethical concerns of exposing participants to misinformation. We still expected to observe higher-order thinking activities as participants will not know whether there will be any errors in the information provided by the videos. The length of the three videos on a topic is set to 30 minutes. The properties of the videos are summarized in Appendix A.

To give participants some note examples to kick-start their video processing and reduce the burden of documentation, we employed extractive summarization algorithms to generate two key highlighting sentences for each video (e.g., as (7) in Figure 1, the note introduces a key position). We utilized a range of techniques, including TextRank, LexRank [34, 77], Latent Semantic Analysis (LSA) [40], and BERT-based language models [87]. We chose the top-rated two-sentence highlights for the videos as identified by most of these algorithms and presented in all study conditions.

5.4 Participants

We recruited 52 participants from a university campus using an attention check that filtered participants, resulting in 13 participants per condition. They are aged 18 to 45 (14 males, 35 females, 3 non-binary). Study conditions and video topics were randomly assigned. Participants assigned to nudge conditions were randomly divided into groups of three. The participants' levels of education ranged from college to PhD. To ensure balanced prior knowledge across groups, we assessed participants' healthcare knowledge using a 7-point Likert scale [57], with a median level of 5.18. Given the potential that people with healthcare-related majors may possess superior knowledge and perceive connections between health-related topics, we checked and found that 1–2 people in each condition self-reported their biology or medical backgrounds, confirming even distribution. There was no group difference between participants' personal relevance with the topic. Their attitudes towards the video topics were also measured, yielding a mean score of 5.69 ($sd = 1.26$), where 1 stands for "strongly anti-[topic]" and 7 stands for "strongly pro-[topic]." Furthermore, as participants' conscious thinking ability [59], education level, and prior attitudes about the video topic largely influenced their higher-order thinking skills and motivation [60], we ran a Kruskal-Wallis test to compare the means in different conditions since the values are not normally distributed. We found no significant differences across study conditions in education level ($\chi^2(3) = 1.45$, $p = 0.69$) and prior attitudes about the video topic ($\chi^2(3) = 0.04$, $p = 0.99$). Similarly, we invited participants to finish the cognitive reflection test [41], and found no significant difference in their System 2 thinking ability [59] across study conditions ($\chi^2(3) = 5.17$, $p = 0.16$).

5.5 Hypotheses

In Study 2, we formulated three hypotheses to explore how peers' notes might influence participants' cognitive engagement, the quality of their thinking, and their attitudes toward the video topic with different thinking nudge presentation tools. Building on the findings from Study 1, we anticipate that the peer interactions in Study 2 will further intensify participants' engagement with heightened presence and awareness of peers' thinking activities. By exchanging thoughts on video topics in real-time and observing how others formulate their notes, participants are likely to deepen their level of engagement when such efforts appear to be normative, and when peer comments are stimulating. Additionally, comment sharing during video viewing will reveal variations in the depth and quality of participants' thinking. Regarding the impact on their attitudes toward the video topic, it is expected that participants' post-experimental attitudes will be influenced by both their initial stance and the attitudes of their peers, if applicable. Research suggests that when participants' attitudes align with their peers or the video content, they are more likely to maintain their stance. Conversely, when there is a discrepancy between their own attitudes and those of their peers or the video, the outcome is less predictable [3, 88]. Pre-study surveys indicated that most participants held pro-[topic] attitudes, which were consistent with the videos, leading us to hypothesize that collaborative work on the videos will reinforce these attitudes. Therefore, our hypotheses are as follows:

- ***H1***: Participants in the nudge conditions will demonstrate a higher level of reflective and critical thinking engagement (***H1a***), produce more thoughtful notes on the canvas or in the document (***H1b***), and exhibit greater agreement with the video topic (***H1c***) compared to those thinking alone.
- ***H2***: Among participants with the same level of nudge availability (whether they all receive and share thoughts or think independently), those using DeepThinkingMap will show a higher level of engagement in thinking (***H2a***), create notes that reflect more complex thinking processes than those using the baseline of Google Docs to collaborate (***H2b***), and display stronger agreement with the video topic (***H2c***) compared to those using a document-based interface.
- ***H3***: Participants with peer thinking nudges and DeepThinkingMap will exhibit significantly higher levels of engagement in thinking (***H3a***), generate more insightful notes (***H3b***), and maintain attitudinal alignment with the video topic (***H3c***) compared to those working independently with Google Docs.

5.6 Data Collection and Analysis

We collected the engagement level of thinking activities via the survey. Also, we measured multiple thinking results, such as attitude and personal sharing, to understand their thinking levels and results of thinking activities. We interviewed participants as well to understand their thinking processes.

5.6.1 Engagement in Thinking. We employed the same survey instrument used in Study 1 (section 4.5.1) [60] to measure participants' engagement in different thinking processes during the study. To facilitate a more refined assessment of thinking engagement, we reviewed the description to match it with the context of the current study design, and converted the survey to a 7-point Likert scale ranging from 1 ("Strongly disagree") to 7 ("Strongly agree"). Cronbach's alpha is 0.868.

5.6.2 Attitude toward Topic. As we design thinking nudge presentations to encourage individuals to reflect on videos and resolve controversies, we measured participants' attitudes toward the video topic, as well as their self-assessed confidence in their attitudes before and after the task. We developed single-question survey items to collect participants' attitudes about the video topic and their levels of confidence about those beliefs using a 7-point Likert scale. Survey questions asked include "To what extent do you agree that [topic] is currently safe?" and "How confident are you about your current opinion?". As discussed above, this work focuses mainly on the *changes* thinking nudges may foster in information workers' engagement and behaviors. There tend to be individual differences in personal attitudes and beliefs, and we are only interested in systematic changes around their attitudes.

5.6.3 Video-related Notes. The video-related personal insights contributed by participants to the canvas or documents (i.e., comments of video reviews) serve as repositories that capture people's externalized thoughts, which may help assess the depth of their thinking processes. First, we identified comments containing personal thoughts that do not directly repeat the video content. We then applied the same coding method from Study 1 to categorize them into three levels of cognitive processes using Bloom's taxonomy [6]: *understanding*, *analysis*, and *evaluation*. The higher the cognitive level revealed from the notes, the greater engagement it is with the reflective and critical thinking activities. We had to restrict the analysis to comments collected from only conditions where DeepThinkingMap was used, as individual contributions can only be clearly segmented and labeled with this interface. In Google Docs, it is technically infeasible to analyze the notes because different participants could have merged multiple points into the same paragraphs. It is infeasible to identify the authorship reliably, especially in a collaborative setting.

Next, we conducted a language analysis to assess language use in the notes. Prior research suggests that specific linguistic features can reveal insights into an individual's thoughts, emotions, and other psychological states [80, 83]. Hence, we utilized the Linguistic Inquiry and Word Count tool (LIWC 22) [20], a dictionary-based text analysis tool, to measure linguistic attributes, including *analytical thinking* and *causation* words. For the control conditions, we scored each participant's notes with LIWC. For the nudge conditions, considering that participants often abstained from noting repeated ideas as learned from our interviews with them, we analyzed the entire set of notes from one group as a single document and assigned a shared LIWC score to each member of the group.

5.6.4 Semi-structured Interview. To further grasp the role of thinking nudges in supporting participants' thinking processes and gathering feedback on DeepThinkingMap, we conducted semi-structured interviews centered around participants' review process, attitude changes, and usage of DeepThinkingMap. After transcribing the interviews, we utilized thematic analysis combining both open and axial coding [76]. In open coding, we pinpointed recurring codes tied to interview topics. These codes were later refined and merged into broader themes during axial coding, ensuring all themes were driven by consensus among the authors [76].

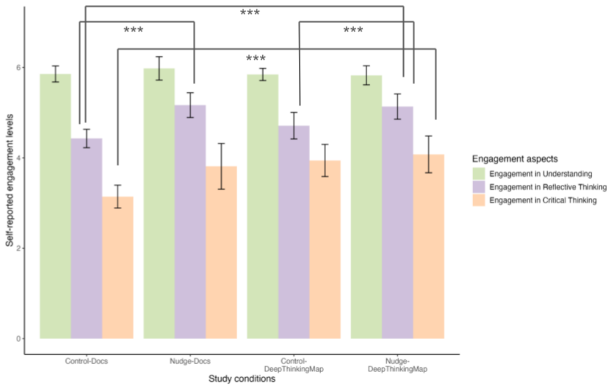


Figure 4: Level of Engagement in Understanding, Reflective Thinking, and Critical Thinking for All Conditions in Study 2. Error Bars Indicate the Standard Error.

5.7 Results

Two-way mixed-effect ANOVAs were conducted to assess the influence of synchronous peer thought exchanges and DeepThinkingMap for different thinking processes. Satterthwaite's method was used to approximate the degrees of freedom in their mixed-effects models, which may lead to non-integer degrees of freedom. The nudge interface (Google Docs vs. DeepThinkingMap) and the availability of thinking nudge are the two independent variables. The video topic was included in the model as a random effect. The demographics, background, education level, prior attitude, and self-reported healthcare background were included as covariate variables. Recognizing the importance of group dynamics, we accounted for the potential influence of diverse educational backgrounds within a group, as differences in education levels can lead to the emergence of leaders or alter group dynamics in nudge conditions. To capture this effect, we considered the contribution of peer participants' educational backgrounds in the 3-person group and added the variance of education levels in each group as another covariate variable (with the value for control conditions assigned as 0).

5.7.1 Engagement in Reflective Thinking. ANOVA results revealed significant differences in both collaborative interfaces and the availability of thinking nudges. As shown in Figure 4 and Table 5 (row 2), individuals or groups using DeepThinkingMap engaged in significantly more reflective thinking than those using Google Docs. With the same interface used, there was a statistically significant effect such that participants in the nudge conditions engaged more deeply in reflective thinking activities compared to thinking-alone participants.

Furthermore, post-hoc Tukey comparisons provided clarity on the pairwise differences between conditions, as in Table 6 (row 1). Regardless of the interfaces used, people in nudge conditions constantly practiced more reflective thinking than those who thought alone. We also found that DeepThinkingMap further amplified this effect.

5.7.2 Engagement in Critical Thinking. For the engagement of critical thinking, we found similar statistical results from ANOVA and post-hoc analysis (Table 5 row 3). Thinking nudges significantly improved the engagement in critical thinking, with participants in nudge conditions delving more profoundly into critical thinking engagement compared to individual participants. Meanwhile, compared to participants using Google Docs, those who utilized DeepThinkingMap exhibited significantly greater critical thinking engagement. Moreover, prior attitude plays a statistically significant role in the level of critical thinking ($\beta = -0.56$, $F(1, 38) = 21.28$, $p < 0.01$, $\omega^2 = 0.33$). Participants who held more positive attitudes toward the topic engaged less in critical thinking since their existing attitudes aligned with the videos.

Following ANOVA analysis, Tukey post-hoc comparisons showed that critical thinking engagement is only significantly different between Nudge-DeepThinkingMap and Control-Docs, as illustrated in Table 6 (row 2). The joint impact of nudging and DeepThinkingMap supports the lift in critical thinking.

5.7.3 Engagement in Understanding. As shown in Figure 4, neither thinking nudges nor DeepThinkingMap exhibited a significant influence on the engagement of understanding. DeepThinkingMap did not significantly impact engagement. The effect of nudging was marginally significant. Statistics are reported in Table 5 row 1. Tukey comparisons further confirmed the lack of significant difference between any two conditions (all $p > 0.3$). The average engagement scores hovered around 6 out of 7 across all conditions, suggesting a consistently high level of understanding engagement, potentially nearing a saturation level for participants.

5.7.4 Video-related Notes. For the content analysis of notes participants left, we report in two dimensions: overall linguistic characteristics and cognitive levels of comments. Figure 5 demonstrates the number of comments and their cognitive levels in DeepThinkingMap-related conditions. A one-way ANOVA model with one independent variable and the same covariant and random variables as other ANOVAs shows that peer-nudged participants created more analysis comments than isolated participants who work in DeepThinkingMap alone. We also found a significant positive impact on evaluation comments when participants were nudged. ANOVA did not show evidence that there were significantly more or less understanding comments in nudge conditions than in control conditions. Detailed statistics are reported in Table 7. The results suggest a positive impact of thinking nudges with DeepThinkingMap on producing critical and reflective thoughts after engaging with these higher-order thinking processes.

For the linguistic results of all notes, including the summary of videos, we connect our findings with prior literature about writing and thinking engagement. The analytical thinking score in LIWC, derived from categories of function words, captures the degree to which people use words that suggest logical and hierarchical thinking patterns [94]. Table 8 and 5 show that notes from nudge conditions had significantly higher analytics thinking scores than those from control conditions. Meanwhile, notes in DeepThinkingMap also show marginally significant odds of higher analytics scores than notes using Google Docs. Within words related to cognitive processes, we went through a full analysis and found that the frequencies of causality keywords (e.g., "how," "because," and

Table 5: Summary of ANOVA Results for Engagement in Higher-order Thinking, Understanding, and Notes by Study Conditions in Study 2.

Measures	Conditions	F stats	p value	ω^2
Engagement in Understanding	Control vs Nudge	$F(1, 44) = 3.48$	0.07	
	Docs vs DeepThinkingMap	$F(1, 44) = 0.32$	0.57	
Engagement in Reflective Thinking	Control vs Nudge	$F(1, 44) = 15.21$	$p < 0.01^{**}$	0.24
	Docs vs DeepThinkingMap	$F(1, 44) = 5.12$	0.03^{*}	0.08
Engagement in Critical Thinking	Control vs Nudge	$F(1, 43.75) = 6.59$	0.01^{*}	0.11
	Docs vs DeepThinkingMap	$F(1, 43) = 6.60$	0.01^{*}	0.11
LIWC Analytics from Notes	Control vs Nudge	$F(1, 44) = 5.16$	0.03^{*}	0.08
	Docs vs DeepThinkingMap	$F(1, 43.16) = 3.76$	0.06	
LIWC Causality from Notes	Control vs Nudge	$F(1, 43) = 0.27$	0.61	
	Docs vs DeepThinkingMap	$F(1, 43) = 8.71$	$p < 0.01^{**}$	0.15

Table 6: Statistically Significant Post-hoc Comparisons on Engagement in Higher-order Thinking in Study 2.

Measures	Group Comparison	Tukey's ω	p value
Engagement in Reflective Thinking	Nudge-Docs - Control-Docs	3.63	$p < 0.01^{**}$
	Nudge-DeepThinkingMap - Control-DeepThinkingMap	3.63	$p < 0.01^{**}$
	Nudge-DeepThinkingMap - Control-Docs	3.98	$p < 0.01^{**}$
Engagement in Critical Thinking	Nudge-DeepThinkingMap - Control-Docs	3.33	$p < 0.01^{**}$

Table 7: Summary of ANOVAs on Comments by Cognitive Categories in DeepThinkingMap Conditions in Study 2.

Measures	F stats	p value	ω^2
Understanding	$F(1, 18.96) = 0.05$	$p = 0.83$	
Analysis	$F(1, 18.87) = 15.40$	$p < 0.01^{**}$	0.41
Evaluation	$F(1, 19) = 15.02$	$p < 0.01^{**}$	0.40

"make") significantly differed between the conditions. Participants using DeepThinkingMap used more causal words in their notes, which may imply more critical thinking, as confirmed from prior studies [80]. However, thinking nudges did not significantly impact the use of casual words. ANOVA results are shared in Table 5 rows 4 and 5.

Table 8: Means and Standard Deviations of LIWC Analysis from the Notes. Standard Deviations are in the Parentheses.

Condition	Analytics	Causes
Control-Docs	72.26 (3.30)	3.84 (0.33)
Nudge-Docs	77.43 (2.56)	3.45 (0.41)
Control-DeepThinkingMap	78.29 (3.07)	4.36 (0.47)
Nudge-DeepThinkingMap	81.36 (1.96)	4.46 (0.35)

5.7.5 Attitude Shift. We are interested in how participants' attitudes and confidence levels change after the experiment. Note that the change in attitudes and confidence levels is a side proxy measure in this study to understand the results of thinking processes.

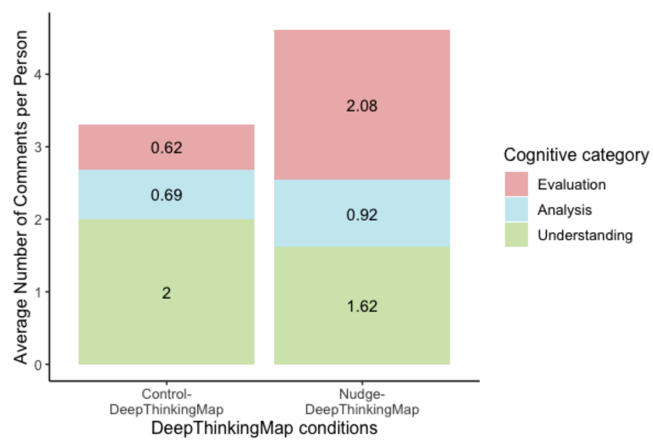


Figure 5: Number of Comments by Cognitive Category in DeepThinkingMap Conditions.

Due to the consistent conclusion for all three videos (pro-[topic]), it is likely that participants would be persuaded as they rarely identified themselves as domain experts. Hence, we focus on the extent to which their beliefs become strongly pro-[topic] and assess their confidence levels. We refined our description regarding attitude polarization, transitioning from a general pro-[topic] or anti-[topic] stance to a more specific pro-[video] or anti-[video] perspective, as the videos are all pro-[topic], in order to mitigate the potential confusion for readers about topic direction.

We conducted two separate mixed-effect ANOVA models, as our attitude measure is in two dimensions: attitudes and self-assessed confidence in their attitudes. As a preliminary test, the correlation between attitude and confidence level is median ($r = 0.56$), implying that they are not strongly connected. For attitude, watching the videos significantly moved participants' attitudes toward pro-[video], aligning with the video materials and the prior attitude from the majority of participants. However, results revealed that neither thinking nudges nor DeepThinkingMap significantly impacted participants' attitudes. As for the confidence of attitudes, watching the videos significantly strengthened participants' confidence in their attitudes. Moreover, ANOVA confirmed the effects of both DeepThinkingMap and thinking nudges on participants' confidence about their attitudes. DeepThinkingMap significantly strengthened attitudinal confidence. Similarly, they also had a positive impact on such confidence. ANOVA statistics are reported in Table 9. From post-hoc Tukey test results reported in Table 10, thinking nudges consistently improved confidence compared to thinking alone. Moreover, we observed that thinking nudges with DeepThinkingMap were more effective in strengthening confidence than solitary work with Google Docs. Figure 6a and 6b illustrate the change in attitudes and their confidence over the topic.

5.7.6 Semi-Structured Interviews. From the interviews, we gained qualitative understandings regarding participants' first-hand experiences with thinking nudges and DeepThinkingMap. Participants in the nudge condition overall appreciated increased accessibility to their peers' video notes. For instance, participants in the Nudge-DeepThinkingMap condition actively integrated their notes with those of others on the canvas. P404 shared that *"if I saw a note from others, I thought my note was important to that as well"*. Moreover, in Study 2, the interactive nudge further enriched the experience. Many reported a heightened sense of social presence in nudge conditions as they may *"feel the social pressure to write notes rather than procrastinate"* (P207). This increased perception of social presence appeared to act as a catalyst, driving participants to contribute their notes and stay productive. Furthermore, they identified multiple advantages conferred by DeepThinkingMap. For example, they reported that the interactive feature for linking different notes on the canvas can help visually *"see how people think about things that are related together"* (P406).

Through the review, participants reported an increase in confidence regarding their perspectives on the video topics. Engaging with peers not only acted as a validation mechanism by reinforcing their viewpoints but also fostered an environment wherein participants felt compelled to actively and critically engage rather than merely fulfilling the study's requirements. For example, P403 mentioned that *"the opinions of my group members seem to go hand in hand with my ideas, and kind of confirmed in my brain that I'm on the right track"*. Meanwhile, participants in nudge conditions also maintained a degree of self-reliance throughout the study. For example, some chose to check the canvas or Google Docs only after completing their video analysis. P417 noted, *Communication and collaboration were very limited as they were not the focus of the group.*

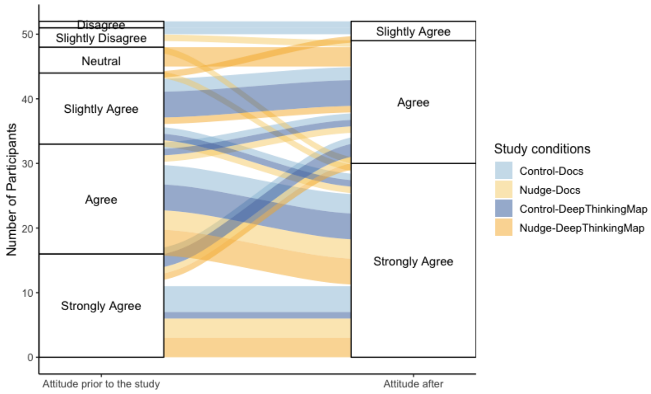
6 Discussion

Higher-order thinking is required for individual and collaborative information work on high-stakes online content but often demands external support. We propose leveraging social-behavioral transparency and peer sharing of thoughts to support higher-order thinking processes. We also comprehensively evaluated DeepThinkingMap's impact on deep thinking. We found that such thinking nudging mechanisms, together with the collaborative concept mapping design, augmented users' critical and reflective thinking engagement (H1a, H2a, H3a supported), led to more thoughtful notes (H1b and H2b supported in Study 2, not supported in Study 1), and increased confidence in participants' correctly held beliefs (H1c, H2c, H3c partially supported). These results are summarized in Table 11. In the subsequent discussion, we discuss the nuanced process of using peer sharing to engage people in higher-order thinking and how that process leads to tangible thoughts and attitudes. We also discuss the potential negative consequences of accessing group opinions in thinking tasks, such as "groupthink" or over-convergence of thoughts due to perceived norms or pressure to conform. Furthermore, we discuss the design implications and delineate the design space for how user-generated content can promote thinking.

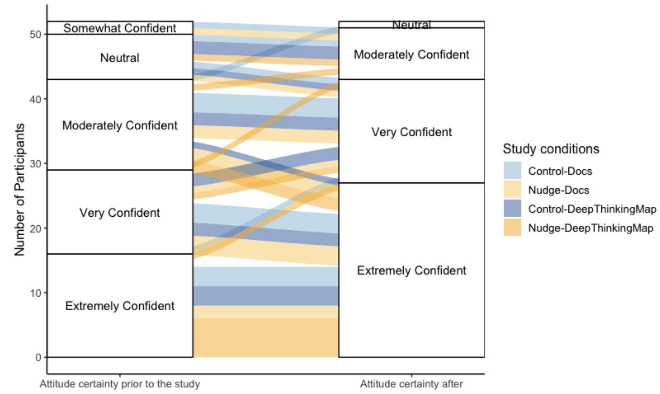
6.1 Nudge to Engage in Thinking and Producing Deep Thoughts

Our studies measured the engagement of higher-order thinking and thoughtful insight production. These are intertwined but distinct measures of associated cognitive activities. In Study 1's non-interactive individual scenario, we did not observe a significant difference in the number of thoughtful comments. However, participants engaged in reflective thinking more actively when the static cues were presented. Their map-editing frequency is positively related to their reflective thinking in the control condition, which aligns with the literature about the positive influence of concept-mapping activities on reflective thinking [54, 97]. Meanwhile, these thinking nudges are fixed and presented as one-offs without interactions, which may limit their influence on deep thought production.

In contrast, thinking nudges were more interactive and dynamic in Study 2, and peers' behavioral cues were more visible. Participants actively merged their notes with peers' notes on the canvas and mentioned their high sense of social presence in nudge conditions. Thus, we observed significant improvement in higher-order thinking engagement and increased numbers of evaluation and analysis comments. Our work identifies the potential of framing collaborations as peer influence through interface design to engage users in reflective thinking and critical thinking, aligning with previous work on utilizing collaboration for reflection [85]. These peer influences could also encourage general users to record and share thoughtful insights in interactive and continuous settings. Additionally, compared to traditional online learning, which typically involves text-based expert-sourced materials, video review tasks involve dynamic, user-generated content that requires temporal and spatial alignment of insights. These findings indicate that designing an efficient deep-thinking nudge requires considering the delivery format and interaction mechanisms of how these thinking-nudge cues are visualized and exchanged to ensure their impact on reflective thinking and written output.



(a) Pre- and post-study Attitude Shift.



(b) Pre- and post-study Attitude Confidence Shift.

Figure 6: Attitude and Attitude Confidence Shift in Study 2.

Table 9: Summary of ANOVAs on Attitude Shift in Study 2.

Measures	Condition and Confounding Variable	F stats	p value	ω^2
Attitude	Control vs Nudge	$F(1, 95.37) = 0.66$	0.41	
	Docs vs DeepThinkingMap	$F(1, 95) = 0.07$	0.79	
	Access Videos	$F(1, 95) = 21.03$	$p < 0.01^{**}$	0.17
Attitude Confidence	Control vs Nudge	$F(1, 95.59) = 8.91$	$p < 0.01^{**}$	0.08
	Docs vs DeepThinkingMap	$F(1, 95.01) = 5.04$	0.03^{*}	0.04
	Access Videos	$F(1, 95) = 14.27$	$p < 0.01^{**}$	0.12

Table 10: Statistically Significant Post-hoc Comparisons on Attitude Shift in Study 2.

Measures	Group Comparison	Tukey's ω	p value
Attitude confidence	Nudge-Docs - Control-Docs	2.96	0.02^{*}
	Nudge-DeepThinkingMap - Control-DeepThinkingMap	2.96	0.02^{*}
	Nudge-DeepThinkingMap - Control-Docs	3.50	$p < 0.01^{**}$

6.2 Influences on Attitudes toward the Topics

To understand the influence of our interventions on attitude, Study 2 tracked participants' attitudes to the video topics before and after the experiment. In the pre-study surveys, the majority of participants held clear attitudes aligned with the video content, and we did not observe any noticeable change in the direction of their attitudes in our study. Meanwhile, participants' confidence in their attitudes increased with peer thinking nudges, where people exchanged takeaways and comments. Such engagement fosters an online environment that encourages people to actively and critically reflect on their attitudes toward the topics. In compliance with research ethics, we only selected videos that included correct information and debunked misleading content. Participants developed stronger confidence in their correctly held attitudes due to these videos and peer comments. Given the participants' enhanced reflective and critical thinking, their increased attitudinal confidence likely reflects careful deliberation, not just self-confirmation.

6.3 Scrutinizing the Threat of Groupthink

Our study proposes that access to peers' understanding and insights promotes deeper reflective and critical thinking. Peers' thoughts serve as a catalyst for more profound, thoughtful engagement with videos. This approach contrasts with the well-documented phenomenon of groupthink, where the desire for group harmony often encourages individuals to conform to the majority opinion, suppressing critical thinking and independent reasoning [56].

Groupthink is more likely to occur in settings characterized by strong group cohesion, small groups, directive leadership, and a lack of diversity in perspectives [64]. While we did not observe directive leadership in our study, participants were selected to ensure diversity along specific dimensions such as culture, gender, or disciplinary backgrounds. There is a possibility that groupthink may emerge, as similar viewpoints might resonate, leading to uncritical acceptance of ideas and discouragement of dissenting opinions.

Table 11: Summary of Findings from Studies 1 and 2.

Impact of	On	Results
Thinking nudge	Engagement in reflective thinking and critical thinking	With static nudge cues, reflective thinking engagement is improved, while critical thinking is not. With interactive nudges, engagement in both higher-order thinking activities is higher. (H1a supported)
	Written thinking results	With static nudge cues, there was no difference in the thoughtfulness of comments. With interactive nudges, there are more comments at reflective and critical thinking levels, and group notes are more analytical. (H1b partially supported)
	Attitude toward the topic	Nudging strengthens attitude confidence, while no significant difference is found for attitude positivity. (H1c partially supported)
DeepThinkingMap	Engagement in reflective thinking and critical thinking	Engagement in both reflective thinking and critical thinking is higher with DeepThinkingMap. (H2a supported)
	Written thinking results	With interactive nudging, notes include more causal words, but are not more analytical. (H2b partially supported)
	Attitude toward the topic	Attitude confidence was stronger with DeepThinkingMap, while no difference is found for attitude positivity. (H2c partially supported)
Interaction Effect	Engagement in reflective thinking and critical thinking	Engagement in reflective thinking and critical thinking is significantly higher when nudging with DeepThinkingMap. (H3a supported)
	Written thinking results	There is no difference when DeepThinkingMap is paired with nudging. (H3b not supported)
	Attitude toward the topic	Attitude confidence was stronger with nudging using DeepThinkingMap. (H3c partially supported)

However, a closer examination of our designs and experimental results suggests the thinking nudge encourages rather than suppresses reflective thinking. As a choice-preserving intervention made available in a non-linear concept mapping space, the social cues may have acted as subtle prompts, allowing participants to integrate their peers' insights at their own pace. This approach preserved cognitive autonomy and facilitated spontaneous higher-order thinking, supported by the results of commenting logs in Study 1 and the results of engagement in reflective thinking and video-related notes in Study 2. Importantly, we did not impose the requirement or format of collaboration in the nudge condition in either study; instead, we fostered a social presence and social access to peer comments that highlighted the individualized processes and outcomes of thinking. As shown in the interview results, some participants in Study 2 nudge conditions reported a substantial degree of perceived independence when analyzing videos. Although none explicitly mentioned using the "Hide Comments" button available on the interface to reduce access to, and thus the influence of, peer comments, some chose to delay their engagement with peer comments until they had watched all the videos and formed their impressions in the video review session. This discretionary behavior suggests that participants felt comfortable controlling social inputs at their preferred pace and timing, which is a novel observation. The observations also imply the thinking-nudge designs could encourage independent thinking while allowing participants to benefit from peer insights without imposing conformity.

Moreover, the heightened level of thinking engagement observed indicates that social loafing was not a contributing factor. Social loafing refers to individuals' social-psychological tendency to exert less effort when working in a group—for instance, due to reduced accountability [91]. Instead, our study found an increased sharing of evaluation comments and cognitive engagement, suggesting that participants were actively reviewing the videos with no sign of loafing.

6.4 Design Implications for Supporting Higher-order Thinking

Prior literature has explored various mechanisms that support thinking-demanding tasks [14, 24]. In this work, we invited non-expert participants to share conceptual understandings and personal thoughts as cues and explored the potential of thinking nudges for deeper thinking in reviewing video materials. Similarly designed interventions from other research, such as social annotations in the educational context, were also shown to benefit students in detailed reading comprehension, learning confidence, and critical thinking [18]. Our findings extend the use of these peer-generated notes to information work beyond traditional education—for example, in contexts where users encounter content from diverse sources with varying credibility. Moreover, our work adds to the literature by showing design possibilities for utilizing average-quality comments from non-expert users with appropriate mediation and representations, in self-supporting and self-sustaining information work. For instance, incomplete content summaries could be utilized as thinking cues for future, non-collaborating users, as in Study 1. These summaries may be designed to support users in attending to specific parts of the information encountered or used to illustrate what other people might think when faced with the same information.

By comparing DeepThinkingMap with conventional shared documents as the mediating tool for information work, Study 2 underscores the necessity of endowing informational representations with interface affordances conducive to prompt higher-order thinking. Non-linear, structured representations of thoughts, such as concept maps, are likely to foster more deliberate and reflective use of peer comments. Our results show that while accessing peer comments can have a general positive effect on thinking, it is crucial to carefully consider how such information is gathered, organized, and displayed for a more sustaining and fulfilling experience of thinking. For instance, the non-linear concept map visualization can be instrumental in communicating peers' thoughts constructively. By allowing users to see connections between ideas in a graphical format, concept maps help users engage more deeply with the material, organizing their thoughts and uncovering insights that might be otherwise overlooked. This approach encourages higher-order thinking by making abstract concepts more reachable and tangible and facilitating a more structured exploration of peer-generated content.

We noted that the use of peer-generated content is not without risks. Unverified and incorrect peer comments can contribute to the dissemination of misinformation, especially when particular individuals dominate the conversation. Redundancy is another concern, as repetitive comments can dilute the impact of a specific idea and overwhelm users. Social annotation studies have also identified these issues [5]. It is essential to design systems that can summarize and curate peer contributions with considerations of information quality and values, ensuring that the thinking-nudge content remains relevant and helpful.

We also noted one technical opportunity lies in the incorporation of AI-generated content. With advancements in large language models (LLMs), generating contextually relevant content to prompt people to think with these models is now possible. Such thinking-nudge cues could be tailored by AI in real-time, adapting to users' states and goals of information work. Despite technological availability, it remains unclear whether using AI-generated comments to encourage deeper thinking is feasible and viable. For example, the behavioral nudging effects may wane when users perceive that the comments presented to them were not produced by human thinking. Ethical concerns may also arise if the use of AI to generate comments is not disclosed to the users. Because AI's cognitive processes differ from humans', its use as a thinking prompt in interactive settings such as Study 2 may yield distinct outcomes warranting further research.

7 Limitations and Future Work

Our experiments have several limitations. First, we used four specific health-related topics, rather than surveying what participants may be interested in beforehand. We tried to mitigate this in the study setup by choosing topics from prevalent social debates like GMOs [74] and popular YouTube videos so that the selected videos simulate videos that participants typically encounter in everyday life. We did statistical tests to make sure the personal relevance of the topic and personal cognitive level were not statistically different; we used participants' education level to mitigate the potential confound of personal relevance to an extent.

We ensured that all the videos shown were credible and consistent with current scientific consensus. This can lead to a lack of observations of how participants react and discuss misinformation within the thinking-nudge content in high-stakes domains. Moreover, our participant sample may not represent diverse demographics, health backgrounds, or attitudes globally. When we recruited participants, the contentious nature of topics like vaccination deterred some from participating. It led to a skewed distribution of prior attitudes in participants, so the findings did not capture much interaction and controversy-resolving between a triad of participants with different group opinion dynamics. Although attitude change is not the sole focus here, further studies could benefit from considering these factors when working on persuasion and thinking nudges for tasks such as misinformation detection and debunking.

Furthermore, our study primarily examined the effects of thinking nudges where peers were unfamiliar with each other. Prior research, however, has suggested that the source of information plays a pivotal role in determining its perceived credibility. For instance, information originating from a reliable friend is often deemed more trustworthy than that from a stranger [46]. As a direction for future research, we aim to study the scenarios when both credible sources and misinformation exist. We also plan to engage participants from varied backgrounds and relationships to comprehensively discern the nuances of social nudges across different demographic groups.

Given the limited observations in a lab setting, the study is not well suited to examine the long-term effect of DeepThinkingMap. We would like to address these limitations further in the future by implementing DeepThinkingMap as a website plugin for daily usage. We plan to conduct longitudinal studies where users may use DeepThinkingMap to freely watch videos of diverse topics and mixed quality and gauge the long-term effects of visualization designs and thinking nudges. Another limitation is that the experiment setting retains the minimum user motivation. We did not divide participants into groups based on their internal motivation levels, nor did we measure motivation as they completed the task. Since motivation and triggers interact—motivation can affect how well a trigger works and triggers can sometimes influence motivation—it is still unclear if our designs acted as triggers that encouraged deeper thinking or increased participants' internal motivation to think beyond the basic study requirements. In our future longitudinal studies, we will also consider the motivational states users have when receiving the thinking nudge or visualization scaffold to have a more nuanced analysis.

8 Conclusion

In this work, we explore the potential of peer sharing as thinking nudges to facilitate critical and reflective thinking when consuming high-stakes video information. We propose to use peers' thinking results as thinking-nudge cues and merge them with a concept-map-based thinking scaffold, DeepThinkingMap. Through two studies, we demonstrate it can effectively engage users in reflective and critical thinking, and the boosting effect of thinking nudge via visualization design. Our findings contribute to a nuanced understanding of the socio-technical-cognitive mechanisms and the design space of peers' sharing visualizations.

Acknowledgments

We thank all anonymous reviewers for their suggestions. The research is supported in part by the University of California, Davis, through a startup fund granted to Hao-Chuan Wang. It is also supported by NSTC grant 113-2221-E-194-030-.

References

[1] 2023. Firebase. "https://firebase.google.com/".

[2] Zhila Aghajari, Eric PS Baumer, and Dominic DiFranzo. 2023. Reviewing Interventions to Address Misinformation: The Need to Expand Our Vision Beyond an Individualistic Focus. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–34.

[3] Dolores Albarracin and Sharon Shavitt. 2018. Attitudes and attitude change. *Annual review of psychology* 69 (2018), 299–327.

[4] Casper Albers and Daniël Lakens. 2018. When power analyses based on pilot data are biased: Inaccurate effect size estimators and follow-up bias. *Journal of experimental social psychology* 74 (2018), 187–195.

[5] Jumana Almahmoud. 2024. *Enhancing Online Collaborative Learning: Designs for Effective In-Situ Discussion and Engagement in Large-Scale Learning Environments.* Ph. D. Dissertation. Massachusetts Institute of Technology.

[6] Lorin W Anderson and David R Krathwohl. 2001. *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.* Longman.

[7] Amid Ayobi, Paul Marshall, and Anna L. Cox. 2020. Trackly: A Customisable and Pictorial Self-Tracking App to Support Agency in Multiple Sclerosis Self-Care. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–15. doi:10.1145/3313831.3376809

[8] Sanchayan Banerjee and Peter John. 2021. Nudge plus: incorporating reflection into behavioral public policy. *Behavioural Public Policy* (2021), 1–16.

[9] Kenneth J Barideaux Jr and Philip I Pavlik Jr. 2023. Enhancing memory recall during video lectures: does the visual display format matter? *Educational Psychology* (2023), 1–20.

[10] Eric P.S. Baumer. 2015. Reflective Informatics: Conceptual Dimensions for Designing Technologies of Reflection. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 585–594. doi:10.1145/2702123.2702234

[11] Eric PS Baumer, Vera Khovanskaya, Mark Matthews, Lindsay Reynolds, Victoria Schwanda Sosik, and Geri Gay. 2014. Reviewing reflection: on the use of reflection in interactive system design. In *Proceedings of the 2014 conference on Designing interactive systems.* 93–102.

[12] Yochai Benkler and Helen Nissenbaum. 2006. Commons-based peer production and virtue. *Journal of political philosophy* 14, 4 (2006).

[13] Marit Bentvelzen, Jasmin Niess, Mikołaj P Woźniak, and Paweł W Woźniak. 2021. The development and validation of the technology-supported reflection inventory. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems.* 1–8.

[14] Marit Bentvelzen, Paweł W. Woźniak, Pia S.F. Herbes, Evropi Stefanidi, and Jasmin Niess. 2022. Revisiting Reflection in HCI: Four Design Resources for Technologies That Support Reflection. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 1, Article 2 (mar 2022), 27 pages. doi:10.1145/3517233

[15] Kristoffer Bergram, Marija Djokovic, Valéry Bezençon, and Adrian Holzer. 2022. The digital landscape of nudging: A systematic literature review of empirical research on digital nudges. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems.* 1–16.

[16] Md Momen Bhuiyan, Michael Horning, Sang Won Lee, and Tanushree Mitra. 2021. NudgeCred: Supporting News Credibility Assessment on Social Media Through Nudges. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–30.

[17] Cristina Bicchieri and Eugen Dimant. 2022. Nudging with care: The risks and benefits of social information. *Public choice* 191, 3-4 (2022), 443–464.

[18] Genevive Bjorn. 2023. The Power of Peer Engagement: Exploring the Effects of Social Collaborative Annotation on Reading Comprehension of Primary Literature. *AI, Computer Science and Robotics Technology* (2023).

[19] Nigel Bosch, Yingbin Zhang, Luc Paquette, Ryan Baker, Jaclyn Ocumpaugh, and Gautam Biswas. 2021. Students' verbalized metacognition during computerized learning. In *Proceedings of the 2021 CHI conference on human factors in computing systems.* 1–12.

[20] Ryan L Boyd, Ashwini Ashokkumar, Sarah Seraj, and James W Pennebaker. 2022. The development and psychometric properties of LIWC-22. *Austin, TX: University of Texas at Austin* (2022), 1–47.

[21] Michael G Brown, James Schiltz, Holly Derry, and Caitlin Holman. 2019. Implementing online personalized social comparison nudges in a web-enabled coaching system. *The Internet and Higher Education* 43 (2019), 100691.

[22] John T Cacioppo and Richard E Petty. 1982. The need for cognition. *Journal of personality and social psychology* 42, 1 (1982), 116.

[23] Alberto J Cañas, Greg Hill, Roger Carff, Niranjan Suri, James Lott, Gloria Gómez, Thomas C Eskridge, Mario Arroyo, and Rodrigo Carvajal. 2004. CmapTools: A knowledge modeling and sharing environment. (2004).

[24] Ana Caraban, Evangelos Karapanos, Daniel Gonçalves, and Pedro Campos. 2019. 23 ways to nudge: A review of technology-mediated nudging in human-computer interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.* 1–15.

[25] Tanya L Chartrand and Rick Van Baaren. 2009. Human mimicry. *Advances in experimental social psychology* 41 (2009), 219–274.

[26] Shiah-Lian Chen, Tienli Liang, Mei-Li Lee, and I-Chen Liao. 2011. Effects of concept map teaching on students' critical thinking and approach to learning and studying. *Journal of Nursing Education* 50, 8 (2011), 466–469.

[27] Robert B Cialdini. 2009. *Influence: Science and practice.* Vol. 4. Pearson education Boston, MA.

[28] Stefano DellaVigna and Elizabeth Linos. 2022. RCTs to scale: Comprehensive evidence from two nudge units. *Econometrica* 90, 1 (2022), 81–116.

[29] Michael Diehl and Wolfgang Stroebe. 1987. Productivity loss in brainstorming groups: Toward the solution of a riddle. *Journal of personality and social psychology* 53, 3 (1987), 497.

[30] Jelle van Dijk, Jirka van der Roest, Remko van der Lugt, and Kees CJ Overbeeke. 2011. NOOT: a tool for sharing moments of reflection during creative meetings. In *Proceedings of the 8th ACM conference on Creativity and cognition.* 157–164.

[31] Vania Dimitrova, Antonija Mitrovic, Alicja Piotrkowicz, Lydia Lau, and Amali Weerasinghe. 2017. Using learning analytics to devise interactive personalised nudges for active video watching. In *Proceedings of the 25th conference on user modeling, adaptation and personalization.* 22–31.

[32] Nicola Diviani, Bas Van den Putte, Corine S Meppelink, and Julia CM van Weert. 2016. Exploring the role of health literacy in the evaluation of online health information: insights from a mixed-methods study. *Patient education and counseling* 99, 6 (2016), 1017–1025.

[33] K Anders Ericsson and Herbert A Simon. 1998. How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking. *Mind, Culture, and Activity* 5, 3 (1998), 178–186.

[34] Günes Erkan and Dragomir R Radev. 2004. Lexrank: Graph-based lexical centrality as salience in text summarization. *Journal of artificial intelligence research* 22 (2004), 457–479.

[35] J St BT Evans, David E Over, and Ken I Manktelow. 1993. Reasoning, decision making and rationality. *Cognition* 49, 1-2 (1993), 165–187.

[36] Jonathan St BT Evans and Keith E Stanovich. 2013. Dual-process theories of higher cognition: Advancing the debate. *Perspectives on psychological science* 8, 3 (2013), 223–241.

[37] Jingchao Fang, Yanhao Wang, Chi-Lan Yang, Ching Liu, and Hao-Chuan Wang. 2022. Understanding the effects of structured note-taking systems for video-based learners in individual and social learning contexts. *Proceedings of the ACM on Human-Computer Interaction* 6, GROUP (2022), 1–21.

[38] Linda Finlay. 2008. Reflecting on 'Reflective practice'. https://api.semanticscholar.org/CorpusID:9580537

[39] Martin Flintham, Christian Karner, Khaled Bachour, Helen Creswick, Neha Gupta, and Stuart Moran. 2018. Falling for fake news: investigating the consumption of news via social media. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems.* 1–10.

[40] Peter W Foltz. 1996. Latent semantic analysis for text-based research. *Behavior Research Methods, Instruments, & Computers* 28 (1996), 197–202.

[41] Shane Frederick. 2005. Cognitive reflection and decision making. *Journal of Economic perspectives* 19, 4 (2005), 25–42.

[42] Brian R Gaines and Mildred LG Shaw. 1995. Collaboration through concept maps. (1995).

[43] Christine Geeng, Savanna Yee, and Franziska Roesner. 2020. Fake News on Facebook and Twitter: Investigating How People (Don't) Investigate. In *Proceedings*

of the 2020 CHI conference on human factors in computing systems. 1–14.
[44] H Reed Geertsen. 2003. Rethinking thinking about higher-level thinking. *Teaching Sociology* (2003), 1–19.
[45] Leonie Gerhards and Christina Gravert. 2016. Because of You I Did Not Give Up - How Peers Affect Perseverance. *ERN: Experimental Economics (Topic)* (2016). https://api.semanticscholar.org/CorpusID:55490136
[46] Kim Giffin. 1967. The contribution of studies of source credibility to a theory of interpersonal trust in the communication process. *Psychological bulletin* 68, 2 (1967), 104.
[47] Henner Gimpel, Sebastian Heger, Julia Kasper, and Ricarda Schäfer. 2020. The Power of Related Articles-Improving Fake News Detection on Social Media Platforms.. In *HICSS*. 1–10.
[48] Murray Glanzer and Anita R Cunitz. 1966. Two storage mechanisms in free recall. *Journal of verbal learning and verbal behavior* 5, 4 (1966), 351–360.
[49] Carlos C Goller, Micah Vandegrift, Will Cross, and Davida S Smyth. 2021. Sharing notes is encouraged: annotating and cocreating with Hypothes. is and Google Docs. *Journal of Microbiology & Biology Education* 22, 1 (2021), 10–1128.
[50] Indranil Goswami and Oleg Urminsky. 2016. When should the ask be a nudge? The effect of default amounts on charitable donations. *Journal of Marketing Research* 53, 5 (2016), 829–846.
[51] Rúben Gouveia, Fábio Pereira, Evangelos Karapanos, Sean A. Munson, and Marc Hassenzahl. 2016. Exploring the Design Space of Glanceable Feedback for Physical Activity Trackers. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Heidelberg, Germany) *(UbiComp '16)*. Association for Computing Machinery, New York, NY, USA, 144–155. doi:10.1145/2971648.2971754
[52] Helen Halbert and Lisa P Nathan. 2015. Designing for discomfort: Supporting critical reflection through interactive tools. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*. 349–360.
[53] Hainora Hamzah, Mohd Isa Hamzah, and Hafizhah Zulkifli. 2022. Systematic Literature review on the elements of metacognition-based Higher Order Thinking Skills (HOTS) teaching and learning modules. *Sustainability* 14, 2 (2022), 813.
[54] Michael J. Hannafin. 1992. Emerging Technologies, ISD, and Learning Environments: Critical Perspectives. *Educational Technology Research and Development* 40, 1 (1992), 49–63. http://www.jstor.org/stable/30219997
[55] Pelle Guldborg Hansen and Andreas Maaløe Jespersen. 2013. Nudge and the manipulation of choice: A framework for the responsible use of the nudge approach to behaviour change in public policy. *European Journal of Risk Regulation* 4, 1 (2013), 3–28.
[56] David Dryden Henningsen, Mary Lynn Miller Henningsen, Jennifer Eden, and Michael G Cruz. 2006. Examining the symptoms of groupthink and retrospective sensemaking. *Small Group Research* 37, 1 (2006), 36–64.
[57] CW Hu. 2013. A New Measure for Health Consciousness: Development of A Health Consciousness Conceptual Model. In *Unpublished paper presented at National Communication Association 99th Annual Convention, Washington, DC (November 2013)*.
[58] Dariusz Jemielniak and Aleksandra Przegalinska. 2020. *Collaborative society*. MIT Press.
[59] Daniel Kahneman. 2011. *Thinking, fast and slow*. Macmillan.
[60] David Kember, Doris YP Leung, Alice Jones, Alice Yuen Loke, Jan McKay, Kit Sinclair, Harrison Tse, Celia Webb, Frances Kam Yuet Wong, Marian Wong, et al. 2000. Development of a questionnaire to measure the level of reflective thinking. *Assessment & evaluation in higher education* 25, 4 (2000), 381–395.
[61] David R Krathwohl. 2002. A revision of Bloom's taxonomy: An overview. *Theory into practice* 41, 4 (2002), 212–218.
[62] Emily R Lai. 2011. Critical thinking: A literature review. *Pearson's Research Reports* 6, 1 (2011), 40–41.
[63] Clive Lawless, Pete Smee, and Tim O'Shea. 1998. Using concept sorting and concept mapping in business and public administration, and in education: an overview. *Educational Research* 40, 2 (1998), 219–235.
[64] Carrie R Leana. 1985. A partial test of Janis' groupthink model: Effects of group cohesiveness and leader behavior on defective decision making. *Journal of management* 11, 1 (1985), 5–18.
[65] Ken Jen Lee, Adrian Davila, Hanlin Cheng, Joslin Goh, Elizabeth Nilsen, and Edith Law. 2023. "We need to do more... I need to do more": Augmenting Digital Media Consumption via Critical Reflection to Increase Compassion and Promote Prosocial Attitudes and Behaviors. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–20.
[66] Arthur Lewis and David Smith. 1993. Defining higher order thinking. *Theory into practice* 32, 3 (1993), 131–137.
[67] Jingxian Liao, Mrinalini Singh, Yi-Ting Hung, Wen-Chieh Lin, and Hao-Chuan Wang. 2023. Conceptcombo: Assisting online video access with concept mapping and social commenting visualizations. In *Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing*. 362–364.
[68] Q Vera Liao and Kush R Varshney. 2021. Human-centered explainable ai (xai): From algorithms to user experiences. *arXiv preprint arXiv:2110.10790* (2021).
[69] Ching Liu, Juho Kim, and Hao-Chuan Wang. 2018. ConceptScape: Collaborative concept mapping for video learning. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–12.
[70] Michael Xieyang Liu, Tongshuang Wu, Tianying Chen, Franklin Mingzhe Li, Aniket Kittur, and Brad A Myers. 2024. Selenite: Scaffolding Online Sensemaking with Comprehensive Overviews Elicited from Large Language Models. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–26.
[71] Blake M. DiCosola III and Gina Neff. 2022. Nudging Behavior Change: Using In-Group and Out-Group Social Comparisons to Encourage Healthier Choices. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–14.
[72] Kapil Chalil Madathil, A Joy Rivera-Rodriguez, Joel S Greenstein, and Anand K Gramopadhye. 2015. Healthcare information on YouTube: a systematic review. *Health informatics journal* 21, 3 (2015), 173–194.
[73] Maximilian Maier, František Bartoš, TD Stanley, David R Shanks, Adam JL Harris, and Eric-Jan Wagenmakers. 2022. No evidence for nudging after adjusting for publication bias. *Proceedings of the National Academy of Sciences* 119, 31 (2022), e2200300119.
[74] Claire Marris. 2001. Public views on GMOs: deconstructing the myths. *EMBO reports* 2, 7 (2001), 545–548.
[75] Richard E Mayer. 2002. Multimedia learning. In *Psychology of learning and motivation*. Vol. 41. Elsevier, 85–139.
[76] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and inter-rater reliability in qualitative research: Norms and guidelines for CSCW and HCI practice. *Proceedings of the ACM on human-computer interaction* 3, CSCW (2019), 1–23.
[77] Rada Mihalcea and Paul Tarau. 2004. Textrank: Bringing order into text. In *Proceedings of the 2004 conference on empirical methods in natural language processing*. 404–411.
[78] Jennifer A Moon. 2013. *Reflection in learning and professional development: Theory and practice*. Routledge.
[79] Joseph D Novak. 2010. *Learning, creating, and using knowledge: Concept maps as facilitative tools in schools and corporations*. Routledge.
[80] Tim O'Riordan, David E Millard, and John Schulz. 2021. Is critical thinking happening? Testing content analysis schemes applied to MOOC discussion forums. *Computer Applications in Engineering Education* 29, 4 (2021), 690–709.
[81] Babette Park, Andreas Korbach, and Roland Brünken. 2020. Does thinking-aloud affect learning, visual information processing and cognitive load when learning with seductive details as expected from self-regulation perspective? *Computers in Human Behavior* 111 (2020), 106411.
[82] Souneil Park, Seungwoo Kang, Sangyoung Chung, and Junehwa Song. 2009. NewsCube: delivering multiple aspects of news to mitigate media bias. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 443–452.
[83] James W Pennebaker, Cindy K Chung, Joey Frazee, Gary M Lavergne, and David I Beaver. 2014. When small words foretell academic success: The case of college admissions essays. *PloS one* 9, 12 (2014), e115844.
[84] Margaret M Plack, Maryanne Driscoll, Maria Marquez, Lynn Cuppernull, Joyce Maring, and Larrie Greenberg. 2007. Assessing reflective writing on a pediatric clerkship by using a modified Bloom's taxonomy. *Ambulatory Pediatrics* 7, 4 (2007), 285–291.
[85] Michael Prilla and Bettina Renner. 2014. Supporting collaborative reflection at work: A comparative case analysis. In *Proceedings of the 2014 ACM International Conference on Supporting Group Work*. 182–193.
[86] Allison K. Rapp, Michael G. Healy, Mary E. Charlton, Jerrod N. Keith, Marcy E. Rosenbaum, and Muneera R. Kapadia. 2016. YouTube is the Most Frequently Used Educational Video Source for Surgical Preparation. *Journal of surgical education* 73 6 (2016), 1072–1076. https://api.semanticscholar.org/CorpusID:2651451
[87] Nils Reimers and Iryna Gurevych. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics. http://arxiv.org/abs/1908.10084
[88] Mark Donald C Reñosa, Jeniffer Landicho, Jonas Wachinger, Sarah L Dalglish, Kate Bärnighausen, Till Bärnighausen, and Shannon A McMahon. 2021. Nudging toward vaccination: a systematic review. *BMJ global health* 6, 9 (2021), e006237.
[89] Rüdiger Schmitt-Beck. 2015. Bandwagon effect. *The international encyclopedia of political communication* (2015), 1–5.
[90] Peter Schultes, Verena Dorner, and Franz Lehner. 2013. Leave a comment! An in-depth analysis of user comments on YouTube. (2013).
[91] Yih-Chearng Shiue, Chao-Min Chiu, and Chen-Chi Chang. 2010. Exploring and mitigating social loafing in online communities. *Computers in Human Behavior* 26, 4 (2010), 768–777.
[92] Ching-Ying Sung, Xun-Yi Huang, Yicong Shen, Fu-Yin Cherng, Wen-Chieh Lin, and Hao-Chuan Wang. 2017. Exploring Online Learners' Interactive Dynamics by Visually Analyzing Their Time-anchored Comments. In *Computer Graphics Forum*, Vol. 36. Wiley Online Library, 145–155.
[93] Cass R Sunstein. 2014. *Why nudge?* Yale university press.

[94] Yla R Tausczik and James W Pennebaker. 2010. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of language and social psychology* 29, 1 (2010), 24–54.
[95] Richard H Thaler and Cass R Sunstein. 2009. *Nudge: Improving decisions about health, wealth, and happiness.* Penguin.
[96] W Ben Towne, Carolyn P Rosé, and James D Herbsleb. 2017. Conflict in Comments: Learning but Lowering Perceptions, With Limits. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 655–666.
[97] William MK Trochim. 1989. An introduction to concept mapping for planning and evaluation. *Evaluation and program planning* 12, 1 (1989), 1–16.
[98] Thiemo Wambsganss, Andreas Janson, and Jan Marco Leimeister. 2022. Enhancing argumentative writing with automated feedback and social comparison nudging. *Computers & Education* 191 (2022), 104644.
[99] Minhong Wang, Bo Cheng, Juanjuan Chen, Neil Mercer, and Paul A. Kirschner. 2017. The use of web-based collaborative concept mapping to support group learning and interaction in an online environment. *Internet High. Educ.* 34 (2017), 28–40. https://api.semanticscholar.org/CorpusID:64831991
[100] Shang Wang, Deniz Sonmez Unal, and Erin Walker. 2019. MindDot: Supporting effective cognitive behaviors in concept map-based learning environments. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–14.
[101] Jessie Wilson, Angie Mandich, and Lilian Magalhães. 2016. Concept mapping: A dynamic, individualized and qualitative method for eliciting meaning. *Qualitative Health Research* 26, 8 (2016), 1151–1161.
[102] Joanna Wolfe. 2008. Annotations and the collaborative digital library: Effects of an aligned annotation interface on student argumentation and reading strategies. *International Journal of Computer-Supported Collaborative Learning* 3 (2008), 141–164.
[103] Qunfang Wu, Yisi Sang, Shan Zhang, and Yun Huang. 2018. Danmaku vs. forum comments: understanding user participation and knowledge sharing in online videos. In *Proceedings of the 2018 ACM International Conference on Supporting Group Work*. 209–218.
[104] Koji Yatani. 2016. Effect sizes and power analysis in hci. *Modern statistical methods for hci* (2016), 87–110.
[105] Amy X Zhang, Lea Verou, and David Karger. 2017. Wikum: Bridging discussion forums and wikis using recursive summarization. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 2082–2096.
[106] Xiaoyu Zhang, Senthil Chandrasegaran, and Kwan-Liu Ma. 2021. Conceptscope: Organizing and visualizing knowledge in documents based on domain ontology. In *Proceedings of the 2021 chi conference on human factors in computing systems*. 1–13.
[107] Arkaitz Zubiaga, Maria Liakata, Rob Procter, Kalina Bontcheva, and Peter Tolmie. 2015. Crowdsourcing the annotation of rumourous conversations in social media. In *Proceedings of the 24th International Conference on World Wide Web*. 347–353.

A Videos used in Study 1 and 2

B Questionnaires

Please answer the questions based on your own experiences during the task on a 7-Likert scale, 1 as strongly disagree, and 7 as strongly agree. The items are adopted from [60].

Understanding

- The videos and task required me to understand the main points mentioned in the video and the general topic.
- To succeed in the task and communicate with others in the task or the future, I needed to comprehend the videos and the general topic.
- I needed to understand the videos in order to perform the task, e.g., communicate with the group or friends, finish the mapping, etc.
- During the experiment, I had to continually think about the videos that I just watched and the general topic.

Reflective Thinking

- I questioned what others may understand and review video content, and tried to think of a better way.
- I thought over what I had been thinking and considered alternatives during the experiment.
- I reflected on video reviews and my actions to see whether I could have improved what I did and said during the experiment.
- I re-appraised my experiences in the task and learned from it, during the experiment.

Critical Thinking

- As a result of this task, I changed my attitude toward this topic.
- The task, including the processes, videos, and notes, has challenged some of my firmly held beliefs.
- As a result of this task, I found there were better ways than my normal way of consuming the videos and reviewing.
- During this experiment, I discovered some faults in what I had previously believed to be right.

Table 12: Meta Information of Videos in Study 1

Title	Video type	View count
Immune System Explained I – Bacteria Infection	Theory explanation	53,195,265
The Truth About the Turmeric in Your Golden Latte	Application	510,020

1. The view count was collected in July 2023.

Table 13: Meta Information of Videos in Study 2

Title	Channel	View count (in millions)	Topic
Are GMOs Good or Bad? Genetic Engineering & Our Food	Kurzgesagt – In a Nutshell	12	GMO
Why are GMOs Bad?	SciShow	3	GMO
The Truth About GMOs	Real Science	0.5	GMO
The Side Effects of Vaccines - How High is the Risk?	Kurzgesagt – In a Nutshell	16	Vaccine
The Science of Anti-Vaccination	SciShow	2.9	Vaccine
COVID-19 Vaccine for 5-to-11-Year-Old Children Explained	Public university channel	0.02	Vaccine

The view count was collected in June 2023. The videos are ordered the same as in the experiment.