Towards Dialogic and On-Demand Metaphors for Interdisciplinary Reading

Matin Yarmand
Computer Science and Engineering
University of California San Diego
La Jolla, California, USA
myarmand@ucsd.edu

Courtney N. Reed
Institute for Digital Technologies
Loughborough University London
London, United Kingdom
c.n.reed@lboro.ac.uk

Udayan Tandon
Computer Science and Engineering
University of California San Diego
La Jolla, California, USA
utandon@ucsd.edu

Eric B. Hekler
CWPHS, FMPH, QI, & Design Lab
University of California San Diego
La Jolla, California, USA
ehekler@health.ucsd.edu

Nadir Weibel
Computer Science and Engineering &
The Design Lab
University of California San Diego
La Jolla, California, USA
weibel@ucsd.edu

April Yi Wang
Computer Science
ETH Zurich
Zurich, Switzerland
april.wang@inf.ethz.ch

Abstract

The interdisciplinary field of Human-Computer Interaction (HCI) thrives on productive engagement with different domains, yet this engagement often breaks due to idiosyncratic writing styles and unfamiliar concepts. Inspired by the dialogic model of abstract metaphors, as well as the potential of Large Language Models (LLMs) to produce on-demand support, we investigate the use of metaphors to facilitate engagement between Science and Technology Studies (STS) and System HCI. Our reflective-style survey with early-career HCI researchers (N=48) reported that limited prior exposure to STS research can hinder perceived openness of the work, and ultimately interest in reading. The survey also revealed that metaphors enhance likelihood to continue reading STS papers, and alternative perspectives can build critical thinking skills to mitigate potential risks of LLM-generated metaphors. We lastly offer a specified model of metaphor exchange (within this generative context) that incorporates alternative perspectives to construct shared understanding in interdisciplinary engagement.

CCS Concepts

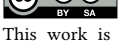
• **Human-centered computing** → **Empirical studies in HCI**.

Keywords

Metaphor Exchange, Large Language Models, Reflective Survey

ACM Reference Format:

Matin Yarmand, Courtney N. Reed, Udayan Tandon, Eric B. Hekler, Nadir Weibel, and April Yi Wang. 2025. Towards Dialogic and On-Demand Metaphors for Interdisciplinary Reading. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26–May 01, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 19 pages. https://doi.org/10.1145/3706598.3713698



CHI '25, Yokohama, Japan

ACM ISBN 979-8-4007-1394-1/25/04
https://doi.org/10.1145/3706598.3713698

1 Introduction

The domain of Human-Computer Interaction (HCI) has historically positioned itself as an interdisciplinary field driven by productive engagement with other disciplines. From its early development, HCI was regarded not as a traditional engineering discipline [41], but a cross-discipline with roots not only in computer science, but also in behavioral and cognitive psychology [23]. Relying on human factors of technology, Don Norman offered a unique role for *human*-centered design: a generalist approach of solving human-related problems by leveraging a breadth of knowledge from fields beyond technology, such as psychology, economics, and business [42].

As the 21st century sees an unprecedented uptick in critical and global-scale challenges, leading HCI figures have put forth a revised interdisciplinary focus for the field. Building on his past seminal definition for the role of design, Norman later revised his stance and proposed *humanity*-centered design to further highlight the key role of social science in informing the next generation of HCI research, in which addressing complex problems of the 21st century is only possible via meaningful engagement with historical, societal, and cultural roots of communities [43]. The unique role of social science in HCI is further echoed in the "Seven Grand Challenges of HCI" [56], in which the main issues facing the world require understanding and planning for cultural shifts, human agency, and humane digital intelligence.

Despite the set vision for HCI, *inter*-disciplinary engagement is often easier said than done, given unique *intra*-disciplinary norms that (to other disciplines) can at best be unfamiliar, and at worst hidden. Disciplines benefit from a uniquely-defined and mutually-agreed set of conventions (in terms of writing style and terminologies) that can be unfamiliar to researchers from other disciplines. For instance, social science literature welcomes papers that deviate from a more structured, engineering-friendly model of IMRaD (Introduction, Methods, Results, and Discussion) [5], a style that is ubiquitous in HCI scholarship. HCI researchers — who are attuned to finding and retrieving information from particular sections of IMRaD-style papers — can struggle when faced with loosely-structured write-ups. Terminologies, which often carry the heavy weight of dense concepts as part of short phrases, can also

pose barriers for entry to other disciplines that are unfamiliar to relevant theories. Social science is no exception, to the extent that extensive effort has gone into creating a dictionary for common jargon [8]. At times, the colloquial use of terminologies can hide their deep-rooted reference to theories from an untrained reader: for instance, Donna Haraway's use of the term "transcendence" [22] stems from feminism literature, yet a lay reader might simply interpret its everyday meaning, resulting in later confusion. Other times, same words could entail different meanings in different disciplines, such as how a term like *modernity* carries a different load in anthropology vs visual arts [55]. While expert input (from within the discipline) can clarify confusions of novice readers, aggregating and providing this support for many existing writing artifacts can be costly and labor-intensive. To engage with new knowledge, those in their early-career or without access to educators must compile and decode this information without support, often reinforcing existing disadvantages.

Many recent computer-supported tools (powered by Large Language Models, or LLMs in short) aim to ease interdisciplinary reading experiences via scalable solutions; yet, it remains unclear to what extent these tools impact overall engagement, especially for a field like social science. Paper Plain [2] generates lay summaries for paragraphs in medical papers, and offers definitions for medical terminologies. Similarly, ScholarPhi [24] assists reading mathematics papers, by generating hyperlinks between unfamiliar symbols and their in-context definitions in the rest of the paper. These tools, while shown to be effective in their respective domains of medicine and mathematics, can lack supporting social science literature that contains dense terminologies in unstructured writing models. Besides, whether these interventions would lead to an increase in meaningful engagement (with the paper) remains an open question. This is especially more important in reading social science papers: anecdotally, a common advice for newcomers in reading social science works is to persevere with reading (and all arising questions and confusions), and ultimately, pieces will fall into place. While confusions (when deliberately scaffolded) can promote learning [17], general education research informs that severely prolonged and unaddressed confusion can lead to boredom, frustration, and eventually disengagement with the learning process [36, 62].

To foster understanding of inter-disciplinary discourse and increase engagement with relevant papers, we investigate using metaphors to bridge between the two domains. Metaphor is a long-standing practice of communicating complex and abstract concepts [33], a foundational tool for exchanging generational wisdom in indigenous communities [67] and facilitating intra-disciplinary education, such as teaching engineering [47] and social science [50]. Unlike the teacher-centered production and transmission of metaphors to students (i.e., receiving agents), contemporary models propose a dialogic exchange of metaphors that empowers the receiving agents to communicate in response and negotiate these metaphoric representations [49]. As argued by Reed et al. [49], this process ultimately leads to a mutual understanding between the two agents (Figure 1).

This paper leverages the dialogic model of metaphor exchange and on-demand properties of LLMs, to investigate bridging between Science and Technology Studies (STS, a sub-discipline of social science) and System HCI (an HCI-subdomain, focused on technology design and development). STS and HCI explore many common topics, yet often via different perspectives; some HCI papers have especially leveraged an STS lens to explore these topics, such as feminist approaches to knowledge creation [3, 10, 60], and intersectionality (i.e., how identity relates to one's surroundings) [52]. As the original model of metaphor exchange was developed via empirical evidence with only teachers (i.e., communicating agents) [49], we particularly aim to explore the receiving agent's perception of the metaphors, and preferred interaction mechanisms to contribute to the mutual representation with the communication agent (i.e., the blue portion of Figure 1). We also generate the metaphors used in this study via LLM, fine-tuned by four key characteristics of effective metaphors [49]. While we do not believe that LLM-generated metaphors serve on-par with carefully-crafted and multi-layered metaphors of people with rich experiential wisdom, LLM can avail on-demand opportunities for metaphor generation and exchange, and, coupled with effective interaction mechanisms for the receiving agent, we envision a productive environment for interdisciplinary reading via metaphor exchange. In short, this paper investigates the following three research questions:

RQ1: *What challenges do HCI scholars face reading STS literature?*
RQ2: *How do on-demand metaphors impact reading STS literature?*
RQ3: *How can HCI researchers interact with metaphors to construct shared understanding with STS?*

This paper conducts a reflective-style survey study (N=48 HCI PhD students from 22 institutions across four continents), and further reports on the challenges faced such as density, writing structure, and perceived non-openness of the STS paper. The metaphors, however, significantly increased the likelihood of reading the provided paper further, depicting promising potential for the use of LLM-generated metaphors in interdisciplinary engagement. The final sketches showcased the HCI participants' desire for flexibility, diversity, and iterative dialogic explanations. The findings of this paper provide important empirical evidence and future guidance for sociotechnical solutions that aim to enhance interdisciplinary engagement in HCI via dialogic and on-demand metaphors.

This study paves the way towards a broader vision of fostering interdisciplinary engagement. While LLMs enable producing on-demand metaphors that are tailored to specific papers and expertise level, they are not void of drawbacks; LLMs can oversimplify, carry biases, and hallucinate. The goal of this work is not to pose LLM as an *end* for understanding papers that are filled with complex lived experiences; rather, LLMs are merely a *mean* to familiarize a novice learner to a sufficient level (ideally, past the ignorant peak of the Dunning-Kruger effect [16]) so they can have a productive conversation with human scholars in the field. The dialogic model of metaphor exchange can especially guide interacting with on-demand metaphors. Our paper mainly addresses one-turn of this turn-taking dialogue in Figure 1, and the rest (i.e., how receiving agents can convey their understanding back to LLMs) is a next step.

2 Related Work

We first provide a brief overview of the rich history of metaphor in HCI, and particularly highlight the dialogic-based models of metaphor exchange. Second, we describe notable LLM-based reading tools that provide on-demand support.

2.1 Metaphor in HCI

"Metaphor" has often been used to refer to the *Desktop Metaphor* [53] and the recreation of real objects within the digital world, or iconic representations of actions [30]. For instance, the desktop itself is meant to be a virtual desk space on which to conduct work, with some research going so far as to attempt to recreate a desk environment in virtual space with life-like physics [1]. Icons such as the trash can represent the act of file deletion, allowing users to drop their files into the trash as they might throw out a document in real life [53].

This iconic digital representation has a tension with interaction and understanding in real-life: Human experiences are being increasingly modulated by technology and we now work with sensory and emotional experiences that are unique to individuals and less simply represented in ubiquitous ways. Indeed, even humans struggle to articulate aspects of experience that exist in wordless knowledge and are rather felt within the body. Contemporary metaphor theory [33, 34] describes how humans communicate conceptual knowledge by mapping a known concept from other life experience [35] onto a target concept we want to convey [4, 33]. For example, we might find the writing of a CHI paper to be *an uphill battle*. Of course, no such battle takes place; the metaphor is *abstract* to the target concept. But, this common idiomatic metaphor allows us to imagine the difficulty in our own experience and ascribe qualities to it (e.g., the metaphor implies the task gets more difficult over time, like going uphill).

Contemporary metaphor theory is of particular interest to HCI because it suggests that we can use our human senses and experiences to abstract concepts into understandable forms [13]. Implementing abstract metaphor into interactive systems would allow the communication of information and experiences where not otherwise possible through an icon or singular representation. However, the challenge lies in creating a metaphor that can be understood by many people, each with differing experiences. Reed et al. [49] offer a model of metaphor communication that describes the process of iterating over a metaphor between the communicating agent and a receiving agent until they are able to negotiate a mutual understanding.

We reference this metaphor model in this paper, along with the four identified features of abstract metaphor, that allow for communication of experience and understanding: metaphor (1) *relies on embodied and lived experiences, rather than pre-existing or domain specific knowledge,* (2) *works independently of language and text-based representations,* (3) *communicates core components and limits unnecessary detail,* and (4) *uses intentional ambiguity to enable individual sense-making* [49].

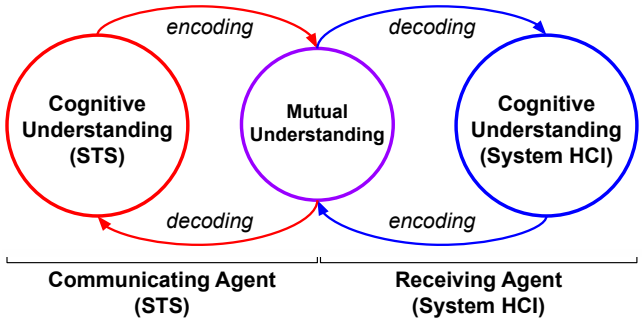


Figure 1: The guiding framework of this research, depicted using the model of metaphor negotiation [49]. Learning via metaphors is through dialogue between the two agents, as well as iterative encoding/decoding with the goal of achieving mutual understanding. This paper specifically considers Science and Technology Studies (i.e., STS) and System HCI.

2.2 LLM-based Reading Tools

The advancement of Large Language Models (LLMs) has led to a wave of tools that leverage the on-demand content-generation capabilities of these models to assist with academic reading. Some tools facilitate domain-general support, such as *CriTrainer* [66] that aims to develop critical reading skills by generating comprehension questions for specific sections. In addition, *CiteRead* [48] enables readers to learn about follow-on works by finding subsequent citing papers, and *CiteSee* [9] augments in-line citations to faster digest and prioritize future explorations. *PaperWeaver* [37] also assists with exploring similar research by further contextualizing short paper alerts, and *Synergi* [31] builds relevant research threads across different papers.

Other tools provide support for specific domains like healthcare and mathematics, with special attention to unique writing norms. For instance, patients often engage with clinical papers that might relate to their own conditions [15], yet are often discouraged given the abundance of unfamiliar terminologies [44]. *Paper Plain* [2] addresses this challenge by generating lay summaries and definitions for the existing jargon. *ScholarPhi* [24] is another example specifically for notation-heavy domains like mathematics and computer science theory, in which symbols and abbreviations are linked to their definition in-context. Other works have especially explored interdisciplinary reading, such as Guo et al. [21] that collected and annotated a large collection of unfamiliar terms. Lastly, *DiscipLink* [68] implements a tri-fold information seeking of orientation, opening, and consolidation.

In this paper, we explore the on-demand capabilities of LLMs to foster interdisciplinary reading. Particularly, we include LLM-generated metaphors as part of our survey study to gauge HCI Readers' perceptiion of understanding unfamiliar STS research. We later discuss the risks of LLMs in the context of metaphors, and offer mitigating guardrails.

3 Methods

We conducted a three-phase, reflective-driven survey to examine the perception and interactions of HCI researchers with metaphoric explanations when reading an STS paper. To ensure mix of expertise and background, we recruited 48 HCI PhD students according to

purposive sampling [19]. Data analysis mainly followed reflexive thematic analysis [7] (for text- and image-based artifacts), as well as statistical and lexical analyses. The Institutional Review Board (IRB) approved the study protocol.

3.1 Survey Design

To answer the three research questions of the study, the survey guided the participants through five sections (questions laid out in Appx. A). *Part 1* addresses RQ1 by gauging challenges when reading an STS text, and *part 2* targets RQ2 via understanding the potential benefits of metaphors. RQ3 shapes the next two parts in terms of how readers digest and interact with these metaphors: *part 3* enables the readers to re-explain the same paragraph (e.g., change metaphor setting or shorten it), and *part 4* prompts reflecting on needs and creating UI techniques (e.g., how to change the metaphor setting or length). To prototype the survey (including the provided metaphors), we ran a pilot study with two HCI researchers who completed the entire survey and thought out loud their impressions.

3.1.1 Part 1: Reading STS. Following the initial survey set-up, the HCI researchers engaged with a short passage of a prominent STS paper: the participants were asked to read the first two paragraphs of Lucy Suchman's Located Accountabilities in Technology Production [57]. Lucy Suchman — a seminal figure in HCI and STS, as well as the recipient of the 2010 SIGCHI lifetime research award — has deep roots in sociology and technology studies; her work is a unique example of interdisciplinary work in the two fields. This paper discusses alternative use and production of technologies informed by recent feminism movements. These interdisciplinary perspectives are crucial for a deeper understanding of the societal and cultural impacts of technology, especially for many system designers and engineers in HCI. In addition, this paper presents unique terminologies in an essay-like format that can be unfamiliar to many HCI researchers.

Following this reading task, the participants answered two questions about how they perceived the provided STS literature (i.e., *RQ1*). First, using Likert-type questions (with 5 scales from *very unlikely* to *very likely*), the survey asked "how likely [the participants are] to continue reading the rest of this paper." Second, an open-ended question inquired about the participants' overall reading experience. The fine prints of the question suggested specific topics that the participants could elaborate in their answer, including writing style and terminologies.

3.1.2 Part 2: Reading STS with metaphors. The second task provided LLM-generated metaphors next to the same STS passage, as depicted in Figure 2.B. To generate these metaphors, we leveraged OpenAI's GPT4-o LLM model [45]. We further tailored the prompts to match the four key features of quality metaphors as noted in Reed et al.'s metaphor-based model of dialogue [49]. We also sanity checked the quality of the metaphors, among a group of experienced HCI and STS scholars. Aligned with the metaphor model, generating an "optimal" metaphor is not the goal; different people can find different metaphors quality given their past experience and expertise. Rather, we ultimately aim to facilitate the dialogue (through the four key features) so that shared understanding is mutually achieved.

To incorporate experiences and perspectives of the two agents (*metaphor feature 1*), we used persona prompting [27], a technique for assigning hypothetical roles and circumstances to the model. Our prompt laid out a scenario in which the LLM was an STS author, explaining STS paragraphs to a System HCI audience. We enabled this role-playing by uploading the top four cited papers of Lucy Suchman, as well as four award-winning CHI papers in the *Blending Interactions* subcommittee (presented on the official website as guiding examples), which centers around technology design and development. We experimented with feeding more papers, and noticed no distinct improvements in the generated metaphors beyond the set of four papers. Additionally, given the potential benefit of multi-modality in metaphor exchange and beyond text-based language (*metaphor feature 2*), we prompted the LLM engine to generate an image for the produced metaphor. Lastly, our prompts included explicit references limiting the level of details and incorporating intentional ambiguity (*metaphor features 3 and 4*), that broadly aim to empower the receiving agents in the metaphor exchange model. To avoid interpreting these metaphors as rigid and direct translations of the original text, these metaphors were titled "alternative explanations" that can invite the receiving agents to openly reflect on the their own perceived usefulness of these metaphors.

Similar to the first part of the survey, the participants completed the same Likert-scale and open-ended question by reflecting on their likelihood to continue reading, as well as reading experience. The Likert-type question posed a hypothetical scenario in which these alternative explanations were readily available for the entire text. The open-ended question (similar to part 1) asked about the overall reading experiences with the fine prints suggesting more granular questions, such as "how the text- and image-based explanation helped/hindered [their] understanding of the original passage" (i.e., *RQ2*).

3.1.3 Part 3: Re-explaining STS. Following the first two reading tasks, the participants attempted to re-explain the STS passage to "help a first time reader understand the original paragraph." This section of the survey aimed to reveal the content of an explanation that might benefit understanding the STS paragraph, without any particular boundary: the question presented a broad space (beyond just the metaphors) by asking the participants to "tweak/modify/overhaul the previous alternative explanation." To minimize the effect of knowledge retention (and instead gauge understanding), this part also provided the original paragraph in-situ of the text box (Fig. 10 in the appendix).

3.1.4 Part 4: Sketching Solutions. The next part of the survey invited the participants to reflect on their overall experience by sketching their "ideal Lucy Suchman Reading Tool." In particular, this step facilitated the space for the participants to reflect and introduce interaction mechanisms that can empower their involvement in the metaphor exchange process (i.e., *RQ3*). Sketching can help elicit *reflective*-driven feedback, in which the participants engage beyond surface level and *reactive*-driven "look and feel" concerns [54]. As a method of reflection, sketching has contributed to deep engagement and idea generation in many areas, such as interior design [59] and medical education [63, 64]. We facilitated three methods of sketching: (1) hand-drawing,

taking a picture, and uploading into the survey, (2) using a free online drawing service, and (3) sketching directly on a widget included in the survey. To help recall past reading experiences (before and after the metaphors), the open ended responses of part 1 and 2 of the survey were displayed back to the participants.

3.1.5 Part 5: Background Information. The last part contained a series of background questions to help characterize the participants. The participating PhD students shared their age, gender, academic institution, years of HCI experience, and backgrounds before starting HCI research. Besides, two Likert-scale questions evaluated the participants' English proficiency skills, informed by common self-rating scales [14]: we gauged the perceived difficulty of listening and writing skills, on seven-option questions starting from "none" to "extreme."

3.2 Participant Recruitment

The distribution of the anonymous survey followed a purposive sampling strategy [19], in which the selection and recruitment of the participants were primarily informed by background and enthusiasm in interdisciplinary research, rather than solely prioritizing diversity (i.e., random sampling) and availability (i.e., convenience sampling). A randomly-selected 10% of the survey respondents received a compensation of US$50 (or equivalent amount in the currency of their residence). To ensure anonymity, we collected the email addresses of the participants on a separate survey that was not linked to the original survey responses.

The participant selection for the survey study pursued a balanced and intentional mix of expertise and experience level. The invited participants all were active PhD students with a history of research contribution at the CHI conference on Human Factors in Computing Systems (i.e., the HCI flagship conference), and many also submitted at other ACM HCI conferences, such as UIST (User Interfaces and System Technologies), UbiComp (Ubiquitous Computing), and CSCW (Computer-Supported and Cooperative Work). This approach welcomed a diverse yet HCI-centered research perspective. Our recruitment strategy further pursued institutional diversity, given that research styles and experiences can heavily rely on the practices and cultural norms of institutions. Meaning, at times we opted to intentionally diverge from recruiting close networks, and instead, target universities and countries that were marginally represented in our pool of participants.

To enhance motivation for completing four sections of an involved survey, we leveraged our interpersonal connections and individually reached out to the HCI PhD students in our network. The authors of this work have a long history of contributing to HCI research, as well as building connections with other researchers from across the globe. Individually reaching out to these candidates — via a variety of communication channels, such as social networks and text messaging — fostered increased trust and enthusiasm in engaging with the survey, especially a longer, reflective-driven survey that involves in-depth reading, writing, and sketching tasks.

3.3 Data Analysis Methods

This section describes the analysis methods for text- and image-based artifacts produced in the survey.

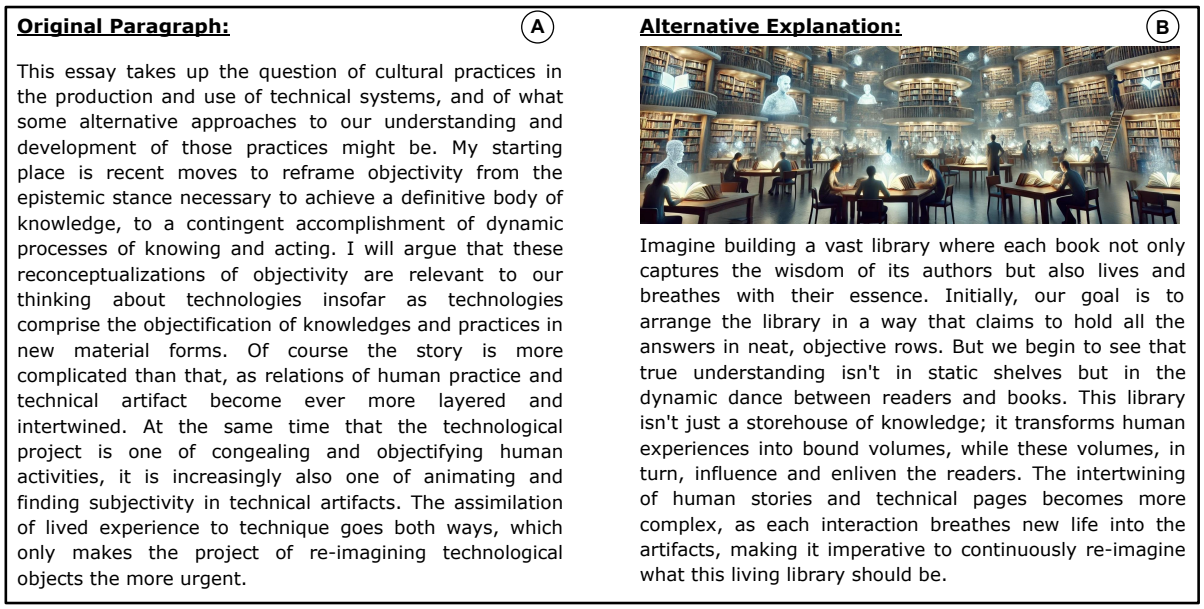


Figure 2: A screenshot of the second reading task in the survey, in which the original paragraph appears on the left, and the (metaphorical) alternative explanation is displayed adjacently on the right. This task further includes the second paragraph of the original passage, and the accompanying alternative explanation describes a large garden with diverse plants in which the role of the gardener is to "nurture, challenge, and transform these diverse perspectives [metaphorically, plants]."

3.3.1 Text-based Thematic Analysis (survey parts 1-2). We analyzed the reading experiences of participants — before and after showing the alternative explanations, as elaborated in parts 1 and 2 of the survey — via a reflexive inductive thematic analysis approach [6, 7]. The first author — an experienced HCI researcher with recent engagement with STS as part of organizing interdisciplinary reading groups — initially built familiarity with the data by actively reading (and re-reading) the responses and identifying relevant content, and later conducted data-driven coding (i.e., open coding). Lastly, pairs of authors iteratively searched and reviewed the codes to create sub-themes, and later themes. The Results section presents these themes.

3.3.2 Statistical Analysis (survey parts 1-2). To statistically measure the differences between the two Likert-scale questions, we employed Wilcoxon test [61]. Wilcoxon is a common statistical method for within-subject analysis with two categorical groups, non-assuming normality. This statistical method fits our Likert-type questions in the first and second parts of the survey, regarding likelihood to continue reading the STS paper.

3.3.3 Lexical Analysis (survey part 3). The attempted re-explanations of the original STS paragraph underwent lexical analysis [32]. Given that many of the participant-generated explanations leveraged different forms of paraphrasing, Keck's taxonomy [32] on paraphrase types provided an analysis lens on the participants' responses. We computed and counted the *unique links* by comparing the produced explanation and the original passage; unique links are individual lexical words and phrases that uniquely appear only once in the original text [32].

3.3.4 Image-based Thematic Analysis (survey part 4). The analysis of the sketches largely followed the same principles laid out in inductive thematic analysis [6], further influenced by suggestions from Tohidi et al. [59]. To build familiarity with the image-based data, two authors printed all sketches and spread them out on a large meeting room table, in which patterns and general motifs quickly became visible. This pair then started re-arranging and open-coding via sticky notes that each displayed an individual code, as well as the relevant sketch ID. These codes mainly represented envisioned features of the sketches, in which we interpreted as interaction mechanisms that can empower the HCI researchers to engage in the communication model of metaphors (i.e., the blue portion of Figure 1). Iterative analysis over the individual codes prompted building and connecting sub-themes, and eventual themes.

4 Results

This section describes the demographics of the survey participants (sec 4.1), as well as the main developed themes grouped by the three research questions of this study. The first subsection describes participant demographics, and the remaining three each address one of the three research questions.

4.1 Demographics of Survey Respondents

Overall, the 48 survey respondents (27 female, 20 male, and 1 non-binary) were affiliated with diverse educational and geographical backgrounds, spanning 22 institutions in four continents. These four continents included North America, Europe, Asia, and Australia, with participants pursuing their PhD degrees from 13, five, three, and one unique institutions, respectively. The USA represented the majority of the survey respondents, with 13 (27% of all participants) from the west coast, and 11 (23%) from the east coast.

The majority of the participants self-reported competent English skills. Many of the participants perceived having advanced reading skills, with 20 (42%) and 14 (29%) respondents selecting none or very little challenges. The remaining 13 (29%) reported some (6; 13%), average (5; 10%), more than average (1; 3%), and much difficulty (1; 3%) with reading. The majority also mentioned high writing skills, including 18 (38%) who self-reported no difficulty and 10 (21%) selecting very little. 8 participants (17%) expressed some difficulty writing English, while the remaining 11 (24%) reported having average (8; 17%) or more than average (3; 7%) difficulty with writing. These results — perhaps as expected due to the dominance of English in research — paint a strong overall language capabilities of the participating PhD researchers.

The participants also represented a diverse age range, HCI experience, and prior academic backgrounds. All participants were in their early adolescence, ranging between 21 and 39, with the average of 28.5 years old. The diversity in age prompted diverse experience levels with HCI research: all participants reported having at least one year of HCI experience, including 18 participants (38%) who engaged with HCI research for three years or less, 26 participants (54%) between four and six years of experience, and the rest (8%) having between 7 and 10 years of experience. Prior to engaging with HCI research, the participants studied diverse disciplines, with the majority of the participants coming from Computer Science backgrounds (21; 44%). Some participants specifically reported sub-fields of Computer Science, such as Machine Learning (2; 4%) and Software Engineering (2; 4%), while some studied other engineering topics, like Electrical Engineering (4; 8%) and Telecommunications Engineering (1; 2%). Outside of the engineering field, the pool of participants further comprised backgrounds in UX Design (4; 8%), Psychology (3; 6%), Philosophy (1; 2%), Architecture (1; 2%), and Biology (1; 2%).

4.2 (RQ1) Challenges of Reading STS by HCI Researchers

This section describes the challenges faced by the participants during the reading part of the survey.

4.2.1 HCI researchers struggled with STS given unfamiliar jargon and writing style, and abstract ideas.

The participants described their experience after reading the two STS paragraphs, highlighting extreme struggles with unfamiliar terminologies and loosely-structured style of writing. They further mentioned the overall denseness of the passage, beyond jargon and writing style.

Many participants struggled to read the provided passage, stemming from many unfamiliar words, especially for non-native English speakers. Some attributed this difficulty to the sheer amount of "*complicated terminology*" (P37), as the paragraphs contained "*many fancy terms and words*" (P13) and "*lots of jargon*" (P19). Additionally, the sub-optimal English skills of some of the

participants further exacerbated the unfamiliarity with the terminologies, such as P47 who shared: "*some words are also uncommon for me as a non-native speaker.*" (P47)

Some participants further mentioned writing style as a core source of difficulty with the provided STS passage, originating from the perceived length and sentence structure. P40, for instance, compared their reading experience to the (more familiar) engineering papers: "*I am more familiar with technical papers, and the sentences are more straightforward/direct comparing to [STS]*" (P40). To explain the source of difficulty with writing style, some participants found the passage to contain "*lengthy, descriptive sentences*" (P2), which contributed to "*losing context midway through them*" (P21). Other participants mentioned sentence and grammatical structure of the passage as key challenges, such as: "*the grammar used in the paper is advanced, which takes me some time to digest*" (P33). P37's difficulty with sentence structure became apparent even on more fundamental levels, as they "*had to look really hard to find the verb in the [second sentence]*" (P37).

Beyond the writing structure and difficult terminologies, the participants further revealed difficulty grasping the overall idea of the STS passage. P21 indicated struggling to understand the overall idea of the author:

> "Despite two coherent sounding paragraphs, I couldn't quite gauge what to expect from the essay. Is it just insights from experience? Or guidelines for future users? Or something else." (P21)

This sentiment was further echoed by P5 who further mentioned high degrees of abstractness: "*the passage is too abstract to make sense of the context in which this opinion is being shared*" (P5). P28 also described a similar experience reading the paper, "*almost like [they] understood individual words but sentences together made no sense*" (P28). Figure 3.A depicts an overview of these challenges via a word cloud of adjectives used to describe reading experience.

4.2.2 *Lack of background knowledge hindered perceived openness and interest.*

The participating HCI PhD students, given their limited backgrounds in social science, felt foreign to and disinterested in the STS passage. Many of the participants acknowledged their lack of knowledge on the topic of discussion in the provided paragraphs. P29 further compared their reading experience with a prior interdisciplinary collaboration: "*I have experienced this before when I collaborate with people in Education research. There's a very different style, which sometimes I learn from, but other times, it feels foreign and confusing*" (P29). P6 mentioned similar feelings of non-openness, as the passage "*seems to rely heavily on the reader's familiarity with the topics it covers. It does not seem to be particularly open to a general audience*" (P6). In addition, prior training (or lack thereof) impacted the interest and motivation of the participants to engage with the provided passage, as reflected: "*I lack background knowledge in feminism and feminist movements, which makes it a bit hard for me to find the inspiration at first glance*" (P33). On the other hand, P44 elaborated that their past interdisciplinary research helped retain motivation and continue engaging with the STS paragraphs:

> "I did not know certain terms, but I could infer them based on the context. However, this is probably due to my familiarity in reading papers that are outside my domain and being okay with not knowing certain terms. However, an undergrad me would have given up." (P44)

4.2.3 *Reading challenges resulted in excessive and unproductive re-reading.*

Multiple rounds of reading the original passage (without gaining distinct understanding) commonly appeared in the participants' responses in the first part of the survey. P16, despite having prior experience with interdisciplinary reading, noted constant re-reading to mitigate lack of perceived understanding:

> "The language is too complex for occasional reading. I've constantly had to stop and return to the beginning of the sentence to find the core stems. Although I've faced a lot with this type of text in my postgraduate education, and I find them extremely valuable (and beautiful from some angle), they are very difficult to comprehend and go through." (P16)

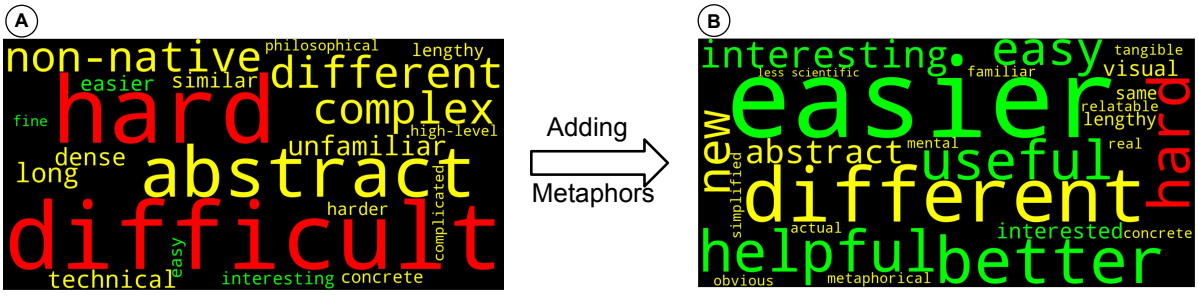


Figure 3: Two word clouds of adjectives that the participants used to describe their reading experience, before and after adding metaphors. The adjectives are coded red (negative), yellow (neutral), and green (positive), based on a common lexicon sentiment analysis [28]. For better readability, we only included words that occurred more than once. These two word clouds depict a stark shift of attitude towards reading experience when having access to metaphors as alternative explanations. The adjectives in (A) relate to the STS passage, while (B) reflects both the STS passage and the provided metaphors.

While some participants justified this re-reading behavior to "*reinforce [their] understanding*" (P32), others admitted that the end-result was not as productive as they had hoped: "*[the STS passage was] hard to follow through. I could not even remember what I had read. I had to read each paragraph multiple times and even now have totally forgotten the second paragraph*" (P39).

4.3 (RQ2) Impact of LLM-generated Metaphors on STS Reading

The following subsections describe the participants' perceptions around the provided metaphors, as well as the impact of these explanations on likelihood to continue reading the STS paper.

4.3.1 Metaphors improved perception of understanding in terms of granular concepts and broader essence.

The participants reported drastic benefits of the provided metaphors, mainly by perceiving better understanding of the original passage. For instance, P29 mentioned the immediate impact of the metaphors in making sense of the STS paragraphs: "*I think the metaphor made the writing a lot easier to understand, and it seemed to click pretty fast*" (P29). Other participants furthered this sentiment by expressing their surprise in how much more they made sense of the original paragraphs: "*I felt like it was a completely different story than the original message - that's how much I didn't understand the original passage*" (P45). Extracting and presenting the adjectives that were used to describe the reading experience (Figure 3–right) visually depicts this change in perceived understanding.

Some participants attributed the helpfulness of the provided metaphors to simpler substitutes for the complex terms and concepts used in the STS passage. Reflecting on the second paragraph (with the metaphor of a garden), P33 linked the STS concepts to granular and lay elements of the metaphors:

> "I find the metaphors used in the alternative explanations very helpful in understanding the concept. For example, instead of trying to think of the process of 'production, reproduction and transformation, etc.', the explanations use the metaphor of growing and planting flowers in a garden to describe the process." (P33)

P7 also benefited from the simplicity of the provided explanation, especially for an English-as-a-second-language researcher, "*because the terms and words used are easier to understand, and after reading the simpler version, reading the original text again made me realize I hadn't understood what the original text was about at all*" (P7).

Some participants underscored connecting with the STS passage on a broader overarching level, without having to understand granular terms and concepts. For example, P32 "*liked how the explanations highlighted the main point of the original paragraphs*" (P32), and P18 referenced understanding the overall gist of the original paragraphs:

> "[The metaphors] helped me better understand the essence of each paragraph and gave me a better context to understand the overall message. [...] It helped through understanding sentences, where I did not understand some words." (P18)

4.3.2 The metaphorical explanations created relatable examples and vivid imagery.

The metaphors further helped the HCI researchers relate to the abstract concepts of the STS passage by using everyday settings and fostering vivid imagination. Many participants praised the metaphors for providing "*tangible*" (P19) and "*more realistic*" (P11) settings which contributed to drawing "*connection from the abstract concepts to real world examples*" (P47). P13 added that these tangible examples increased the appeal to the provided explanation: "*creating real-life analogies with everyday objects made the explanation more interesting*" (P13). In addition, many participants elaborated that engaging with the metaphors enabled creating vivid imagery of the abstract concepts in the original passage: "*the alternative explanations are more accessible than the original text, because they are based on concrete imagery which is easier to reason about than abstract concepts*" (P35). P9 echoed this sentiment and highlighted the ability to "*visualize the concepts better*" (P9) and P20 expressed that "*vividly describing the theory into pictures [made them] interested in reading further*" (P20). Lastly, some participants reported feeling immersed in the visualization:

> "I feel the text improved a lot to be more attractive, by starting with 'imagine' I feel myself more immersed in the situation that is described in the text." (P25)

4.3.3 Metaphors enhanced likelihood to continue reading the STS paper.

Following the first reading task, the participants expressed narrow intent to continue reading the rest of the paper. As shown in Figure 4-top, 60.5% of the participants replied with *very unlikely* and *unlikely* to continue engaging with the STS literature. Some participants attributed their interest for reading the paper to the difficulty they faced in the task: *"as a novice to the field, I found this difficult to read and thus uninteresting"* (P2). P44 also elaborated on their reading struggles and lack of motivation to continue reading, further describing their experience as *"tedious"* (P44): "*the writing style is a bit complicated to understand, makes the experience a bit tedious, and does not generate the necessary engagement that keeps you wanting to read it*" (P44).

Having access to metaphors (similar to ones provided in the second reading task of the survey) significantly increased the likelihood to continue engaging with the STS passage. Figure 4-bottom revealed a stark shift, in which 56.3% of the participants responded with *likely* or *very likely* to continue reading the STS paper. A Wilcoxon test showed that the differences between the two groups were statistically significant ($Z = 83.5$; $p < .001$). Spearman rank-order correlation [26] revealed that English reading skills did not impact likelihood of reading for both without ($p = .25$) and with metaphors ($p = .21$). P40 elaborated that the provided metaphors boosted their interest in reading, resulting in more continued engagement with the original STS passage:

> "Normally, if the text/reading is not what I am interested in or not what I can understand, I will give up easily. With the alternative explanations, it does trigger my interest to read the original text more carefully and trying harder to understand it." (P40)

4.3.4 The images mostly helped via increasing appeal, setting the context, and reinforcing details.

Many participants highlighted that including images complemented the provided metaphors by boosting interest, readying context, and reinforcing details. First, some participants praised the images for the added interest in engaging with both the original and metaphoric explanations. For instance, P28 mentioned that "*the images kept [the metaphors] engaging*" (P28), and P33 shared that "*these images make the [original] passage more appealing and help illustrate the concept vividly*" (P33). Second, the images helped set the scene not only for reading content that is quite different than the original passage, but also for the participants' imagination. In particular, "*the images at the start, created a mental model of what to expect [from the metaphoric explanation]*" (P13), and "*helped with quickly choosing a setting for [their] imagination*" (P19). Third, the granular elements of the provided images helped grasping details of the metaphors, for instance "*some aspects of the images (size of the library, diversity of flowers in the garden) helped reinforce understanding [the alternative explanation]*" (P32). P44 — who acknowledged the benefits of the images, yet sought more details — further highlighted the importance of granular visual elements in the provided images: "*using a picture might be useful to clarify the analogy better in certain cases, if the picture is descriptive enough*" (P44).

Some participants, however, raised indifferent and concerned perceptions around the provided images. Few of the responses questioned the benefit of the visual representation altogether, such as: "*I am not sure if the image was totally necessary, but I personally appreciated it*" (P29). Others felt that the style of the image raised doubts about its helpfulness: "*perhaps because of the GenAI-aesthetic, I ended up not paying too much attention [to the images]*" (P17). Other responses expressed concerns regarding the images, such as P1 who felt the images were "*distracting*" (P1) and "*forced*" (P1).

4.3.5 Some participants valued the alternative viewpoints, others raised lack of trust.

Some participants praised the metaphor and perceived these alternative explanations as truly *alternative*, in which these metaphoric texts were possible (and not the only) way to interpret the original text. For instance, P17 compared their reading experience to a live tutoring session in which the alternative nature of the explanation could uniquely resonate to different perspectives:

> "I liked the explanations, especially with complex passages, they felt like I had someone there who tried to explain to me in a different way that might be easier for me, like a teacher. It offered a different perspective, like a possible interpretation." (P17)

P36 further acknowledged the alternative text to differ significantly than their own interpretation, and found this difference in perception especially interesting:

> "I found [the alternative text] interesting to read because they interpreted what I read completely differently. [...] I think this is also what the text by Suchman is about: having different angles on things can provide a discourse around objectivity." (P36)

Despite the original intention of labeling the metaphors as *alternative* explanations, some participants expected a tight link between the two texts, which further led to doubting the provided text. For instance, P3 questioned whether the alternative text can "*always project the exact thought*" (P3) of the main paper, and P40 expressed uncertainty about how the alternative text was "*translated from the original text*" (P40). Ultimately, these doubts might have resulted in "*how much [the participants] can trust the explanations*" (P46). In addition, some participants raised the importance of knowing the source of the explanation, in which "*it should be obvious that it is a separate text, not written by the author*" (P17). Similarly, P29 further highlighted peer input to build trust:

> "I think I'd need some metric to help me know if I can 'trust' this translation, e.g., if my peers said 'oh this is a great alternate, more digestible text that communicates the important points of the other paper,' then I might be more inclined to trust that it would help my understanding." (P29)

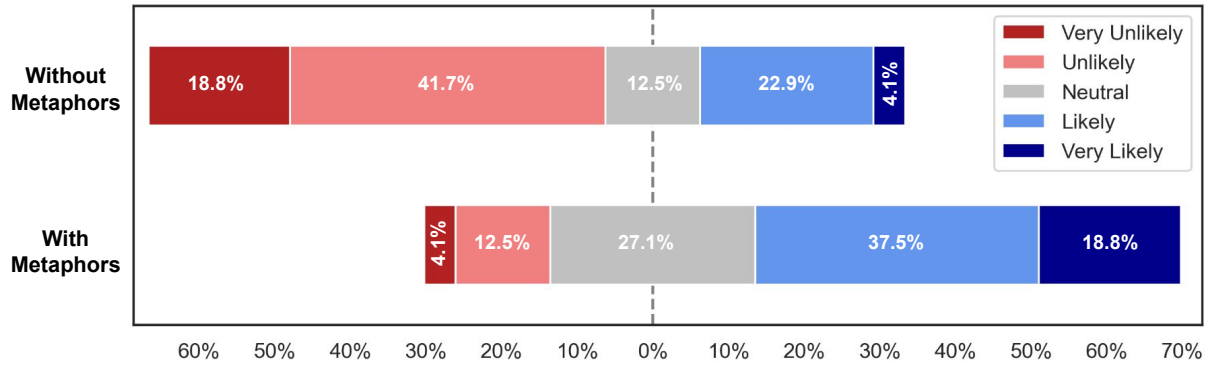


Figure 4: Two diverging bar charts describing the participant's answer to *"how likely are you to continue reading the rest of the paper"*, suggesting the potential of metaphors to enhance engagement with dense and unfamiliar STS literature for HCI researchers. A Wilcoxon test showed statistically significant differences across the two groups (Z=83.5, p<.001).

4.4 (RQ3) Interaction Mechanisms of HCI Researchers in Metaphor Exchange

The sketches facilitated a space for the HCI participants to reflect on the provided metaphors and feel empowered to design interaction mechanisms for engaging with these alternative explanations (i.e., the blue arrows in Figure 1). Table 1 depicts an overview of these themes, grouped into six categories (referred to as "C" for the rest of this section)[1]. We developed these categories via an inductive analytical lens (i.e., bottom-up) with the guiding question on how HCI researchers would interact with these alternative text to construct shared understanding with STS (RQ3). C1–C5 highlight unique properties of *on-demand* metaphors, and C6 highlights a *dialogic* space.

4.4.1 C1) Explanation Granularity.

The sketches showcased a need for diverse granularity for the provided explanations that only referenced one entire paragraph. For instance, S29 (the sketch, S, by P29) in Figure 5 displays mapping the terms in both the original passage and alternative text. In part 2 of the survey, P29 elaborated on having access to explanations for smaller lexical components (e.g., words and phrases) that can provide a tighter link between the two text:

> "I would be interested in some thing that would help equate the two texts at some point in the simplified version. E.g., a 'flower' in this metaphor refers to some jargon word in the original text." (P29)

Other participants preferred to engage with more holistic descriptions (e.g., on a document level), such as S10 that includes a dedicated button to generate an explanation for the entire text. Coarse-level explanations can also mitigate the additional readings needed with the produced metaphors, as some participants raised concerns about the length of the alternative explanation: "*I would be more likely to keep reading the paper if the alternative explanations were shorter. Otherwise it would probably double my reading time*" (P32).

4.4.2 C2) Crowd Perspectives.

Many sketches highlighted the benefit of crowd perspectives in enhancing the provided explanation, mainly in terms of other readers' sense-making of the metaphor and the original passage, other researchers' use of the introduced concepts in their own papers, and other theorists' construction and interpretation of these ideas which might at times differ from the original author. For instance, S33 (Figure 5–right) facilitates a comprehensive space for readers to engage and share interpretations, not only by direct exchange of definitions, but also sharing recent news that might pertain to the provided ideas (and further contribute to grounding these principles into concrete, everyday phenomena). Besides, S31 envisions a space for the authors who cited the STS paper, and in particular the ways that these concepts are further applied, and developed, in different contexts. Lastly, S36 highlights the need for influential theorists and thinkers who might have a different lens of interpreting concepts: in addition to Lucy Suchman (the original author of the STS paper), S36 enables "switching characters" with Plato, Bell Hooks (American theorist), and Marshall McLuhan

[1]We included all the sketches in the supplementary material and as part of an interactive webpage.

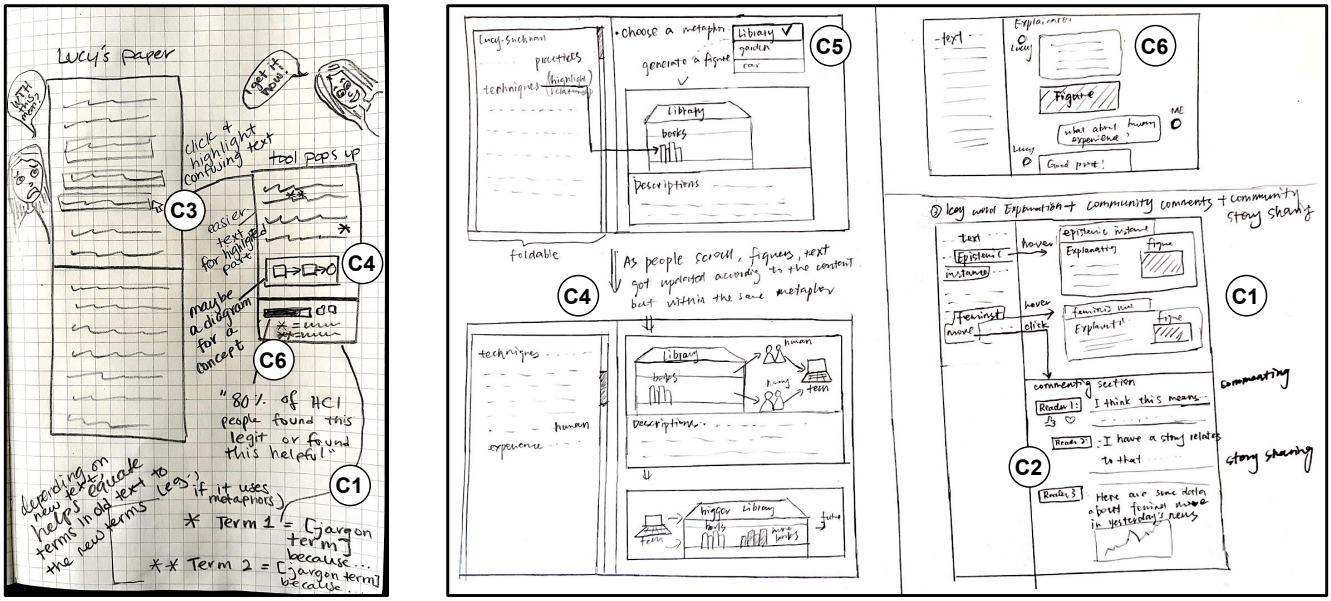


Figure 5: Two hand-drawn sketches by P29 (left) and P33 (right). The sketches aimed to enable the HCI researchers to reflect on interaction mechanisms to better internalize and manipulate the provided metaphors, resulting in a mutual representation with STS. C1–C6 depict examples for the six categories of themes developed from the produced sketches, elaborated in Table 1.

(Canadian philosopher) to communicate their thinking about particular parts of the paper.

4.4.3 C3) Paraphrasing Methods.

The participant sketches portrayed different ways of building mutual understanding with the STS paper, besides providing explanations in the form of metaphors. Summarizing and simplifying were two common methods in the sketches, such as providing a short (1-2 sentence) summary for an 8-10 sentence chunk of the original passage, as displayed in S32. In addition to a summary tab, S37 includes an Explain-Like-I'm-5 (known as ELI5, an internet slang to ask for simple explanations). S37 further demonstrates this feature via an example for a technical biology text that reads "ocean liquid appears to the human sensorial system as an electromagnetic wave"; the ELI5 panel (adjacently on the right) says "the water is blue because of light beams getting their red color absorbed by (lots of) water." S21 also suggests an in-place simplification, in which simple translations (e.g., "closely related") replace the complex jargon of the paper (e.g., "intertwined"), yet the readers can see the original terms by hovering over these replacements.

The generated re-explanations (in part 4 of the survey) further showcased the benefit of more direct interpretations of the STS passage (in the form of summarized and simplified text). Most of the participants created paraphrases (36 out of 44; 81.8%) via a mix of these two techniques. Further analysis of these produced paraphrased content (using Keck's taxonomy [32]) revealed that most re-explanations revised the original paragraph moderately and minimally: as shown in Figure 6, 20 responses (55.6%) fell under *moderate revision* (i.e., 1-19% similarity), while *minimal revision* (i.e., 20-49% similarity) contained 13 paraphrases (36.1%). Eight participants still attempted to re-explain the passage via metaphors, as displayed in Appx. B): four participants used similar setting as the provided alternative explanation (i.e., *library*), while

Category	Theme	Example	Sketch
C1) Explanation Granularity	Word-level	Providing explanations for specialized terms in the paper (S29).	S16; S17; S29; S33; S35; S40; S47
	Sentence-level	Highlighting a sentence produces an explanation only for the selected portion of the text (S17).	S12; S17; S38
	Paragraph-level	Displaying the alternative explanation via a separate button after every paragraph (S45).	S17; S30; S37; S45
	Document-level	Converting the entire text into a simplified explanation using a button at the end (S10).	S5; S10
C2) Crowd Perspectives	Readers' Comprehension	Providing interpretations, personal anecdotes, and relevant news from other readers (S33).	S3; S33; S44; S46
	Authors' Application	Engaging with other authors in terms of how they "cited/phrased/interpreted" the original text (S31).	S31; S37
	Theorists' Interpretation	Adding additional perspectives from well-known philosophers like Plato and McLuhan (S36).	S36
C3) Paraphrasing Methods	Summarizing	Generating a "1-2 sentence summary for an 8-10 sentence paragraph" of the original text (S32).	S5; S12; S16; S21; S32; S37
	Simplifying	Offering simple explanations such as an "explanation for a kid" (S30).	S7; S21; S29; S30; S37; S39; S42
C4) Accompanying Visuals	Setting Context	Displaying a series of images at the start to set the "overall theme" of the paper (S8).	S8
	Concept Map	Generating "concept maps" (S26).	S26; S29
	Comic Style	Updating the figures and text of the initial image according to the content of the original text (S33).	S33
C5) Re-generating Explanations	Re-paraphrasing	Including a "regenerate button" for the provided paraphrased statements (S27).	S5; S27; S34
	Changing Setting	Enabling to switch the setting of the explanation, such as from library to garden (S47).	S33; S47
	Recursive Explanations	Selecting a piece of the provided explanation recursively to produce more explanations (S12).	S6; S12
C6) Iterative Discourse	Chatting with AI	Incorporating a chatbot fine-tuned with the persona of the author, providing a space for questions like "what about human experience" (S33).	S18; S27; S33; S42
	Collaborative Commenting	Providing a collaborative space to showcase how "others are making sense of the content" (S44).	S3; S29; S33; S44

Table 1: Overview of the themes extracted from the sketches, grouped by six categories. This table depicts key interaction mechanisms that can help HCI researchers better construct mutual representation with the STS paper.

the remaining four attempted making connections using *cooking*, *human senses*, and *the river system*. Four participants did not provide re-explanations, and further highlighted the difficulty they faced with the original reading: "*oh sorry I do not think I understand this original paragraph, so can't really help*" (P40). To examine whether the English writing skills influenced how the participants re-explained the original paragraph (i.e., four categories of paraphrasing, metaphor, and no explanation), we ran a Spearman rank-order correlation test [26] and found no significant effect (p = .1).

4.4.4 C4) Accompanying Visuals.

Following the generally positive perceptions around multi-modality of the metaphors (Sec. 4.3.4), some sketches provided concrete purposes and styles for the provided images, such as setting context, mapping concepts, and continually building on the initial images like a comic book. S8, for instance, displays a reading interface for the book, Invisible Women [46] that includes two pictures at the top to set the scene: one, on the left, that shows a graffiti with repeated "GRL PWR"s, and the right image that shows a smart watch. The two images together aim to prompt the readers to gendered data bias, an overarching theme of the book. As shown in Figure 5–left, S29 shows the potential relationship between the presented concepts as part of a diagram. Lastly, S33 introduces a continuity between the images across the STS paper. As displayed in Figure 5–right, scrolling through the viewer of the original passage updates the initial image for the metaphor to incorporate the newly introduced ideas: for instance, the original image that might show a library with only books, can later display their authors and the relationship between the two entities.

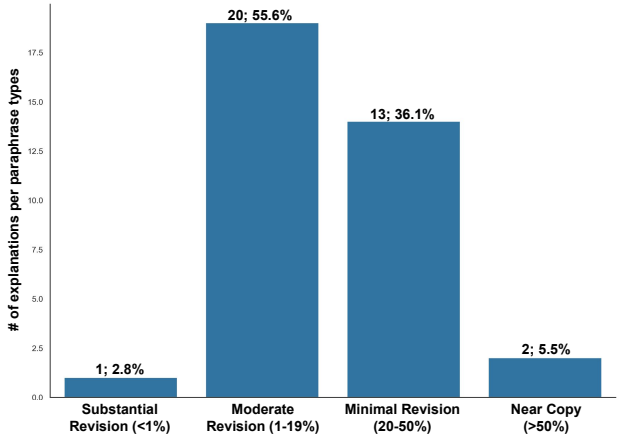


Figure 6: The distribution of the participants' explanations categorized by the degree of paraphrasing, according to Keck's taxonomy [32]. Most of the metaphoric explanations fall within *Moderate Revision* (20 out of 36; 55.6%) and *Minimal Revision* (13; 36.1%). Out of the remaining participants, eight respondents provided metaphoric explanations, as presented in Appx. B, and the last four offered no explanation, citing the difficulty they had with the original passage.

4.4.5 C5) Re-generating Explanations.

The sketches exemplified a key aspect of the participants' desire for agency, not only in terms of activating support, but also re-generating the provided explanations to better support their unique needs. Some participants highlighted having control over receiving support, by challenging the way that the survey readily presented the explanations next to the STS passage. For instance, P3 desired more agency over generating support: "*maybe the alternative text appears on demand if the user needs more help in understanding the text? Showing it always by default can be distracting*" (P3). P17 echoed this sentiment, and in particular valued the initial struggle of sense-making:

> "I would probably not want to see [the explanation] next to the text right away. [...] I believe that, sometimes there is value in first attempting to understand and interpret a text by yourself. Even if it is hard." (P17)

One type of mechanism showed re-generating the provided explanations, such as S34 with designated buttons to produce a new paraphrase for the provided explanation. Other sketches enabled changing the setting of the metaphors, such as Figure 5–right that offers switching the overall theme of the metaphors from library, to garden or a car. Lastly, some sketches expressed the need for depth-first re-generation of support, especially for readers who might seek further explanations into the provided explanation. S12 showcases this recursive functionality, in which each time the system generates a high-level answer, readers can dig deeper into these high-level answers.

4.4.6 C6) Iterative Discourse.

Besides having access to diverse perspectives (as noted in C2), some sketches portrayed mechanisms for iterative and alternating exchange of perspectives, demonstrated via chat bots and comment sections. AI-enabled chat bots appeared in a number of sketches to showcase the potential need for discourse in a flexible format, not strictly defined by the provided explanations. S33 includes an example of a chat bot, in which the reader can engage in open-ended conversations with a model that is specifically trained on the artifact of the main author, as shown in Figure 5–right. Besides, S44 envisions a collaborative space for sense-making, in which readers of the STS reading can engage in a commenting section and successively share their unique perspectives around the paper.

5 Discussion

This section draws out unique benefits of dialogic and on-demand metaphors in the context of interdisciplinary reading, and further lays out important guardrails to ensure faithful alignment with the original papers.

5.1 Dialogic metaphors from different perspectives can draw out inter-disciplinary subtleties

The results highlighted the role of alternative explanations in providing perspectives that might uniquely resonate with readers, and some participants desired more transparency around the source and process of metaphor generation, which ultimately impacted

their trust in the explanations (Sec. 4.3.5). Specifically, the sketches revealed the value of how other readers make sense of the same passage, how other researchers apply these topics in their own domains, and how theorists define and interpret similar concepts (Sec. 4.4.2). Lastly, the participants valued engaging with iterative and flexible discourse to further construct meaning, in the form of chat bots and collaborative commenting sections (Sec. 4.4.6).

Providing diverse perspectives can draw out subtleties of the STS readings, especially to HCI readers with varying background and intentions. As noted in the contemporary model of metaphor exchange [49] (Fig. 1), no singular system of understanding exists, and instead, it is only via a dialogic engagement that mutual spaces of understanding can be constructed. In addition to this dialogic engagement, facilitating an inter-web of perspectives, and especially others' mutual understandings, has the potential to more distinctly uncover the existing subtleties. This can especially benefit interdisciplinary reading by availing these subtleties to researchers who plan on partial engagement with other disciplines. For instance, in the case of this research, HCI readers might have different intentions of engaging with STS literature: while some fully embrace the new discipline, others might have to resort to more limited engagement due to constraints of time and resources. This inter-web of perspectives can especially benefit the latter group, in which it more distinctly reveals the nuances of the STS work in a way that, otherwise (and without in-depth engagement with the field) might go unnoticed. Figure 7 portrays this concept in a specified version of the original metaphor model in the context of interdisciplinary reading.

5.2 On-demand metaphors can foster engagement via intervening at severe confusion points

The HCI researchers who severely struggled to connect with the STS passage (Sec. 4.2), by large found the provided metaphors a welcomed addition to their reading experience, not only in terms of better perceived understanding (Sec. 4.3.1), but also the ability to form concrete and vivid imagery of the concepts (Sec. 4.3.2). The participants, however, noted the importance of autonomy in explanation generation and re-generation, especially to enable initial grappling with complex ideas (Sec. 4.4.5). Lastly, the Likert-scale questions revealed a significantly higher likelihood to continue engaging with the provided STS paper (Sec. 4.3.3), given access to alternative explanations.

The provided metaphors (that can be produced on-demand via LLMs) improved perception of understanding by addressing core confusions, contributing to likelihood for continued engagement. Perception of understanding (via self-reported measures), while different than real understanding (in the form of learning gains), can importantly keep the reader engaged throughout the activity, as shown by prior works in the space of self-directed learning [65]. On-demand access to metaphors (especially during severe confusion) can especially lower the entry barrier by setting concrete examples and imagery early on. On-demand systems of communication emphasize the active role of information seekers to initiate support when deemed appropriate [11], and further borrow ideas from just-in-time models of feedback exchange [51] and productive struggle [25]. In the case of interdisciplinary reading, we envision the alternative explanations to be readily available (to address severe struggles at the point of disengagement), yet not instantly accessible to allow space for constructing mutual understanding out of the vagueness and complexity of the paper.

We believe that the advancement of LLMs has enabled a unique opportunity to leverage this technology for on-demand metaphors, via generating content that can uniquely tailor to individual readers (given their lived experiences and background knowledge), as well as on a broad scale for various pieces of literature. In the next section, however, we raise important risks with LLMs and metaphors, and ways of mitigating these risks.

5.3 Developing critical thinking skills can mitigate risks of LLM-generated metaphors

The use of metaphors to foster interdisciplinary engagement can pose some key risks, especially with respect to a dialogic and on-demand exchange model enabled by LLMs. As reported in Sec. 4.2 regarding limited (and in many cases, lack of) initial understanding, many participants perceived the provided metaphor as "correct", likely due to their constrained background knowledge that might have hindered fact- and sanity-checking. In

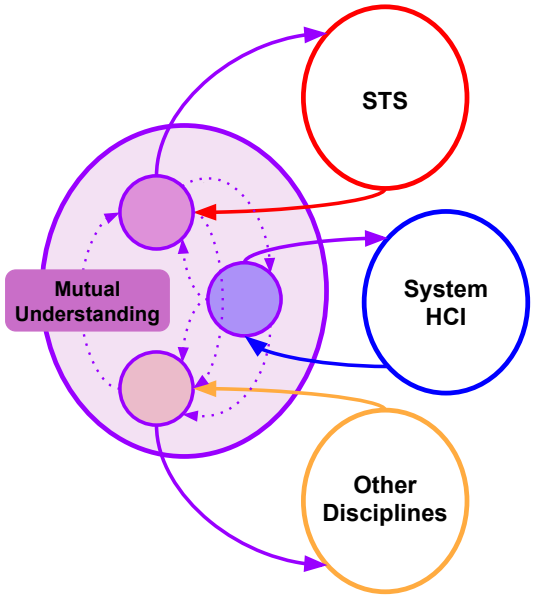


Figure 7: A specified model of abstract metaphor exchange, within the context of dialogic and on-demand metaphors for interdisciplinary reading. Disciplinary knowledge is negotiated to reach mutual understanding (purple) of a target field. Perspectives might originate in a particular field, for instance from STS in this study (red) and domain specific knowledge must be negotiated with understanding from System HCI (blue) or other disciplines (yellow). The resulting mutual understanding is shaped by these disciplines, which are themselves further influenced in turn.

addition, metaphors have a strong suggestive power [58], and mis- and dis-representations of metaphors can have harmful impact, not only on individual ways of knowing, but also collective ways of acting; metaphor misuses have led to continued gendered stereotypes [38], and deliberate exploitation of metaphors has fumed violent wars, such as the pursuit of capitalism in the early 19th century by utilizing Darwin's "survival of the fittest" [12]. Despite the on-demand and dialogic properties of LLMs (which we have built on in this work), these imperfect models are void of lived experiences and perspectives, and are subject to oversimplification, or worse, "hallucination": generation of realistic-sounding but unfaithful content [29]. Overly simplified content can also lead to lack of productive struggle [25] and diminish meaningful learning. As such, while we see promising potential for LLMs to foster interdisciplinary engagement, it is important to acknowledge long-term risks and offer guardrails.

Added perspectives and discourse from the crowd, especially ones close to the field, can mitigate risks via developing critical thinking skills. The participant sketches laid out specific methods of involving other perspectives in sense-making, such as the distinct benefits of other readers, authors, and thinkers (Sec. 4.4.2), as well as iterative discourse (Sec. 4.4.6). As demonstrated in Figure 7, these perspectives can (and perhaps, should) come from people who are active contributors in the field (i.e., STS): researchers with unique and robust understanding of concepts and ideas, beyond select papers. Besides the role of these perspectives in constructing own mutual understanding, these alternative viewpoints can further develop *critical thinking* skills, defined as reflective and reasonable skills of thinking in order to shape own belief system [18]. In particular, engaging with real-world cases and discourse are two core elements of building critical thinking skills [40], both incorporated in the dialogic model of metaphor exchange with diverse perspectives (Sec. 5.1).

5.4 Paraphrasing methods can complement metaphor-based explanations

As evident in Sec. 4.4.3, the participants used varying methods of providing re-explanations for the original paragraph: while many survey respondents employed paraphrasing (via summarization and simplification techniques), some created metaphoric explanations, not only using the provided library examples, but also connecting with other phenomena like human senses and cooking. The participant sketches further reflected this duality of needs for HCI researchers: for instance, as captured in Fig. 5, S29 features a mechanism to produce simplified explanations, while S33 heavily delves into using and manipulating abstract metaphors. The results of this work suggest that learners need different types of explanations when engaging with unfamiliar works.

This diverseness of support needs (beyond solely metaphors) might suggest differing educational goals of researchers engaged in interdisciplinary reading, according to Bloom's taxonomy [20]. Recall of information and direct explanations (i.e., paraphrasing) might relate more closely to *remembering* and *understanding* stages of learning, the lower-level building blocks of learning. Meanwhile, using metaphors can entail *applying* and *analyzing* knowledge by making connections between concepts and real-world phenomena. For instance, a total novice might at first find metaphorical explanations too abstract for sense-making, and might instead benefit from simplified paraphrases. This could also be a back-and-forth process. As such, with the ultimate goal of staying engaged with unfamiliar literature and improving learning, a mix of both metaphoric and paraphrase-based explanations can especially benefit interdisciplinary reading. Future works in this space can further reveal the dynamics between metaphors and paraphrases for long-term learning.

6 Limitations and Future Work

Despite the novel empirical evidence offered in this work, some limitations exist, particularly regarding study design. We used survey to investigate use of metaphors for interdisciplinary reading. Collecting familiarity to STS research (beyond training background) can provide a more well-rounded demographic information. In addition, while surveying helped reach diverse participants widely, more in-depth and flexible methods of data collection (e.g., interviews) can provide deeper insight into the provided responses. We also relied on self-reported data to gauge *perception* of understanding without explicitly measuring *true* understanding or learning gains. While measuring understanding — without reader's unique knowledge constructions and author's unique creative intentions — is challenging (if not impossible), we believe that future work can employ discourse-based learning evaluations [39]. The survey also used one of Lucy Suchman's papers; material from other prominent STS scholars can further complement findings of this study. Lastly, the survey lacked a control condition to evaluate the metaphors; comparing understanding and knowledge retention of reading STS with vs. without metaphors is a natural next step for this research.

7 Conclusion

This paper investigates leveraging dialogic and on-demand metaphors to foster interdisciplinary reading between System HCI and Science and Technology Studies (STS). While engaging with STS is a cornerstone research direction of HCI, this engagement often breaks due to unfamiliar terminologies, writing style, and abstract concepts, as reported in this study. We later highlight that metaphors enhance perception of understanding, and ultimately, desire to continue reading. We lastly discuss the value of different perspectives to draw out subtleties of the new domain, and further argue that while LLM-generated metaphors can address severe confusions on-demand, it is key to acknowledge and mitigate risks via critical thinking skills. This paper provides fundamental empirical evidence in using metaphors for academic reading; when carefully designed for, metaphors can provide important pathways for sociotechnical systems to foster academic reading, especially in a field like HCI that thrives on interdisciplinary engagement.

Acknowledgments

We thank the Dieter Schwarz Stiftung Foundation and the Swiss Confederation for generously funding this search. We also sincerely thank all the participants of the study who offered their time and wisdom on completing the survey.

References

[1] Anand Agarawala and Ravin Balakrishnan. 2006. Keepin' It Real: Pushing the Desktop Metaphor with Physics, Piles and the Pen. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montréal, Québec, Canada) *(CHI '06)*. Association for Computing Machinery, New York, NY, USA, 1283–1292. doi:10.1145/1124772.1124965

[2] Tal August, Lucy Lu Wang, Jonathan Bragg, Marti A. Hearst, Andrew Head, and Kyle Lo. 2023. Paper Plain: Making Medical Research Papers Approachable to Healthcare Consumers with Natural Language Processing. *ACM Trans. Comput.-Hum. Interact.* 30, 5, Article 74 (Sept. 2023), 38 pages. doi:10.1145/3589955

[3] Shaowen Bardzell. 2010. Feminist HCI: taking stock and outlining an agenda for design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 1301–1310. doi:10.1145/1753326.1753521

[4] Pippin Barr, Robert Biddle, and James Noble. 2002. A Taxonomy of User-Interface Metaphors. In *Proceedings of the SIGCHI-NZ Symposium on Computer-Human Interaction* (Hamilton, New Zealand) *(CHINZ '02)*. Association for Computing Machinery, New York, NY, USA, 25–30. doi:10.1145/2181216.2181221

[5] Lluís Codina Bonilla. 2022. The IMRaD model of scientific articles: what is and how can be applied in the humanities and social sciences? *Hipertext. net: Revista Académica sobre Documentación Digital y Comunicación Interactiva* 24 (2022), 6. doi:10.31009/hipertext.net.2022.i24.01

[6] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. doi:10.1191/1478088706qp063oa

[7] Virginia Braun and Victoria Clarke. 2021. One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology* 18, 3 (2021), 328–352. doi:10.1080/14780887.2020.1769238

[8] Craig Calhoun. 2002. *Dictionary of the Social Sciences.* Oxford University Press.

[9] Joseph Chee Chang, Amy X. Zhang, Jonathan Bragg, Andrew Head, Kyle Lo, Doug Downey, and Daniel S. Weld. 2023. CiteSee: Augmenting Citations in Scientific Papers with Persistent and Personalized Historical Context. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 737, 15 pages. doi:10.1145/3544548.3580847

[10] Ko-Le Chen, Rachel Clarke, Teresa Almeida, Matthew Wood, and David S. Kirk. 2017. Situated Dissemination through an HCI Workplace. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 2078–2090. doi:10.1145/3025453.3025696

[11] Yan Chen, Steve Oney, and Walter S. Lasecki. 2016. Towards Providing On-Demand Expert Support for Software Developers. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 3192–3203. doi:10.1145/2858036.2858512

[12] Gregory Claeys. 2000. The "Survival of the Fittest" and the Origins of Social Darwinism. *Journal of the History of Ideas* 61, 2 (2000), 223–240. doi:10.2307/3654026

[13] Claudia Daudén Roquet and Corina Sas. 2021. Interoceptive Interaction: An Embodied Metaphor Inspired Approach to Designing for Meditation. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 265, 17 pages. doi:10.1145/3411764.3445137

[14] Fred Davidson and Grant Henning. 1985. A self-rating scale of English difficulty: Rasch scalar analysis of items and rating categories. *Language Testing* 2, 2 (Dec. 1985), 164–179. doi:10.1177/026553228500200205

[15] Suzanne Day, Stuart Rennie, Danyang Luo, and Joseph D. Tucker. 2020. Open to the public: paywalls and the public rationale for open access medical research publishing. *Research Involvement and Engagement* 6, 1 (Feb. 2020). doi:10.1186/s40900-020-0182-y

[16] David Dunning. 2011. The Dunning–Kruger Effect. In *Advances in Experimental Social Psychology*. Elsevier, 247–296. doi:10.1016/b978-0-12-385522-0.00005-6

[17] Sidney D'Mello, Blair Lehman, Reinhard Pekrun, and Art Graesser. 2014. Confusion can be beneficial for learning. *Learning and Instruction* 29 (Feb. 2014), 153–170. doi:10.1016/j.learninstruc.2012.05.003

[18] Robert H Ennis. 1985. A logical basis for measuring critical thinking skills. *Educational Leadership* 43, 2 (1985), 44–48.

[19] Ilker Etikan, Sulaiman Abubakar Musa, Rukayya Sunusi Alkassim, et al. 2016. Comparison of Convenience Sampling and Purposive Sampling. *American Journal of Theoretical and Applied Statistics* 5, 1 (2016), 1. doi:10.11648/j.ajtas.20160501.11

[20] Mary Forehand. 2010. Bloom's Taxonomy. *Emerging Perspectives on Learning, Teaching, and Technology* 41, 4 (2010), 47–56.

[21] Yue Guo, Joseph Chee Chang, Maria Antoniak, Erin Bransom, Trevor Cohen, Lucy Lu Wang, and Tal August. 2023. Personalized Jargon Identification for Enhanced Interdisciplinary Communication. *arXiv preprint arXiv:2311.09481* (2023).

[22] Donna Haraway. 1988. Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. In *Feminist Studies*. Vol. 14. JSTOR, 575. doi:10.2307/3178066

[23] H. Rex Hartson. 1998. Human-computer interaction: interdisciplinary roots and trends. *J. Syst. Softw.* 43, 2 (Nov. 1998), 103–118. doi:10.1016/S0164-1212(98)10026-2

[24] Andrew Head, Kyle Lo, Dongyeop Kang, Raymond Fok, Sam Skjonsberg, Daniel S. Weld, and Marti A. Hearst. 2021. Augmenting Scientific Papers with Just-in-Time, Position-Sensitive Definitions of Terms and Symbols. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 413, 18 pages. doi:10.1145/3411764.3445648

[25] James Hiebert and Douglas A Grouws. 2007. The effects of classroom mathematics teaching on students' learning. *Second Handbook of Research on Mathematics Teaching and Learning* 1, 1 (2007), 371–404.

[26] Myles Hollander, Douglas A. Wolfe, and Eric Chicken. 2015. *Nonparametric Statistical Methods.* Wiley. doi:10.1002/9781119196037

[27] Tiancheng Hu and Nigel Collier. 2024. Quantifying the Persona Effect in LLM Simulations. *arXiv preprint arXiv:2402.10811* (2024).

[28] C. Hutto and Eric Gilbert. 2014. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media* 8, 1, 216–225. doi:10.1609/icwsm.v8i1.14550

[29] Ziwei Ji, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Ye Jin Bang, Andrea Madotto, and Pascale Fung. 2023. Survey of Hallucination in Natural Language Generation. *ACM Comput. Surv.* 55, 12, Article 248 (March 2023), 38 pages. doi:10.1145/3571730

[30] Jeff A. Johnson, Teresa L. Roberts, William Verplank, David C. Smith, Charles H. Irby, Marian Beard, and Kevin Mackey. 1989. The Xerox Star: a retrospective. *Computer (Long Beach Calif.)* 22, 9 (Sept. 1989), 11–26.

[31] Hyeonsu B Kang, Tongshuang Wu, Joseph Chee Chang, and Aniket Kittur. 2023. Synergi: A Mixed-Initiative System for Scholarly Synthesis and Sensemaking. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology* (San Francisco, CA, USA) *(UIST '23)*. Association for Computing Machinery, New York, NY, USA, Article 43, 19 pages. doi:10.1145/3586183.3606759

[32] Casey Keck. 2006. The use of paraphrase in summary writing: A comparison of L1 and L2 writers. *Journal of Second Language Writing* 15, 4 (Dec. 2006), 261–278. doi:10.1016/j.jslw.2006.09.006

[33] George Lakoff. 1993. The contemporary theory of metaphor. In *Metaphor and Thought*. Cambridge University Press, 202–251. doi:10.1017/cbo9781139173865.013

[34] George Lakoff and Mark Johnson. 1980. *Metaphors We Live By.* University of Chicago Press, Chicago.

[35] George Lakoff and Mark Johnson. 1999. The embodied mind. In *Philosophy in the flesh: Embodied mind and its challenges to Western thought*. New York, Basic Books, 16–41.

[36] Reed W. Larson and Maryse H Richards. 1991. Boredom in the Middle School Years: Blaming Schools versus Blaming Students. *American Journal of Education* 99, 4 (1991), 418–443.

[37] Yoonjoo Lee, Hyeonsu B Kang, Matt Latzke, Juho Kim, Jonathan Bragg, Joseph Chee Chang, and Pao Siangliulue. 2024. PaperWeaver: Enriching Topical Paper Alerts by Contextualizing Recommended Papers with User-collected Papers. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 19, 19 pages. doi:10.1145/3613904.3642196

[38] Emily Martin. 1991. The Egg and the Sperm: How Science Has Constructed a Romance Based on Stereotypical Male-Female Roles. *Signs: Journal of Women in Culture and Society* 16, 3 (April 1991), 485–501. doi:10.1086/494680

[39] Neil Mercer. 2004. Sociocultural discourse analysis: analysing classroom talk as a social mode of thinking. *Journal of Applied Linguistics* 1, 2 (May 2004), 137–168. doi:10.1558/japl.2004.1.2.137

[40] Barak Miri, Ben-Chaim David, and Zoller Uri. 2007. Purposely Teaching for the Promotion of Higher-order Thinking Skills: A Case of Critical Thinking. *Research in Science Education* 37, 4 (Jan. 2007), 353–369. doi:10.1007/s11165-006-9029-2

[41] William Newman. 1994. A preliminary analysis of the products of HCI research, using pro forma abstracts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, Massachusetts, USA) *(CHI '94)*. Association for Computing Machinery, New York, NY, USA, 278–284. doi:10.1145/191666.191766

[42] Don Norman. 2013. *The Design of Everyday Things, Revised and Expanded Edition.* Basic Books.

[43] Don Norman. 2023. *Design for a Better World Meaningful, Sustainable, Humanity Centered.* MIT Press.

[44] Emily Nunn and Stephen Pinfield. 2014. Lay summaries of open access journal articles: engaging with the general public on medical research. *Learned Publishing* 27, 3 (July 2014), 173–184. doi:10.1087/20140303

[45] OpenAI. 2024. Hello gpt-4o. https://openai.com/index/hello-gpt-4o

[46] Caroline Criado Perez. 2019. *Invisible Women: Data Bias in a World Designed for Men.* Abrams Press.

[47] Nicole Pitterson, Natasha Perova-Mello, and Ruth Streveler. 2018. Engineering Students' Use of Analogies and Metaphors: Implications for Educators. *International Journal of Engineering Education* 35, 1 (2018), 2–14.

[48] Napol Rachatasumrit, Jonathan Bragg, Amy X. Zhang, and Daniel S Weld. 2022. CiteRead: Integrating Localized Citation Contexts into Scientific Paper Reading. In *Proceedings of the 27th International Conference on Intelligent User Interfaces* (Helsinki, Finland) *(IUI '22)*. Association for Computing Machinery, New York, NY, USA, 707–719. doi:10.1145/3490099.3511162

[49] Courtney N. Reed, Paul Strohmeier, and Andrew P. McPherson. 2023. Negotiating Experience and Communicating Information Through Abstract Metaphor. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 185, 16 pages. doi:10.1145/3544548.3580700

[50] Debra C. Rosenthal. 1982. Metaphors, models, and analogies in social science and public policy. *Political Behavior* 4, 3 (1982), 283–301. doi:10.1007/bf00990109

[51] Hillol Sarker, Moushumi Sharmin, Amin Ahsan Ali, Md. Mahbubur Rahman, Rummana Bari, Syed Monowar Hossain, and Santosh Kumar. 2014. Assessing the availability of users to engage in just-in-time intervention in the natural environment. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Seattle, Washington) *(UbiComp '14)*. Association for Computing Machinery, New York, NY, USA, 909–920. doi:10.1145/2632048.2636082

[52] Ari Schlesinger, W. Keith Edwards, and Rebecca E. Grinter. 2017. Intersectional HCI: Engaging Identity through Gender, Race, and Class. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 5412–5427. doi:10.1145/3025453.3025766

[53] Ben Shneiderman and Pattie Maes. 1997. Direct manipulation vs. interface agents. *Interactions* 4, 6 (Nov. 1997), 42–61. doi:10.1145/267505.267514

[54] Hilary Smith, Geraldine Fitzpatrick, and Yvonne Rogers. 2004. Eliciting reactive and reflective feedback for a social communication tool: a multi-session approach. In *Proceedings of the 5th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques* (Cambridge, MA, USA) *(DIS '04)*. Association for Computing Machinery, New York, NY, USA, 39–48. doi:10.1145/1013115.1013123

[55] Don Spanner. 2001. Border Crossings: Understanding the Cultural and Informational Dilemmas of Interdisciplinary Scholars. *The Journal of Academic Librarianship* 27, 5 (Sept. 2001), 352–360. doi:10.1016/s0099-1333(01)00220-8

[56] Constantine Stephanidis, Gavriel Salvendy, Margherita Antona, Jessie Y. C. Chen, Jianming Dong, Vincent G. Duffy, Xiaowen Fang, et al. 2019. Seven HCI Grand Challenges. *International Journal of Human–Computer Interaction* 35, 14 (July 2019), 1229–1269. doi:10.1080/10447318.2019.1619259

[57] Lucy Suchman. 2002. Located accountabilities in technology production. *Scandinavian Journal of Information Systems* 14, 2 (2002), 7.

[58] Richard Swedberg. 2020. Using Metaphors in Sociology: Pitfalls and Potentials. *The American Sociologist* 51, 2 (May 2020), 240–257. doi:10.1007/s12108-020-09443-3

[59] Maryam Tohidi, William Buxton, Ronald Baecker, and Abigail Sellen. 2006. User sketches: a quick, inexpensive, and effective way to elicit more reflective user feedback. In *Proceedings of the 4th Nordic Conference on Human-Computer Interaction: Changing Roles* (Oslo, Norway) *(NordiCHI '06)*. Association for Computing Machinery, New York, NY, USA, 105–114. doi:10.1145/1182475.1182487

[60] Kelly B. Wagman and Lisa Parks. 2021. Beyond the Command: Feminist STS Research and Critical Issues for the Design of Social Machines. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 101 (April 2021), 20 pages. doi:10.1145/3449175

[61] Robert F Woolson. 2007. Wilcoxon signed-rank test. *Wiley Encyclopedia of Clinical Trials* (2007), 1–3.

[62] Diyi Yang, Miaomiao Wen, Iris Howley, Robert Kraut, and Carolyn Rose. 2015. Exploring the Effect of Confusion in Discussion Forums of Massive Open Online Courses. In *Proceedings of the Second (2015) ACM Conference on Learning @ Scale* (Vancouver, BC, Canada) *(L@S '15)*. Association for Computing Machinery, New York, NY, USA, 121–130. doi:10.1145/2724660.2724677

[63] Matin Yarmand, Chen Chen, Kexin Cheng, James Murphy, and Nadir Weibel. 2024. "I'd be watching him contour till 10 o'clock at night": Understanding Tensions between Teaching Methods and Learning Needs in Healthcare Apprenticeship. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 457, 19 pages. doi:10.1145/3613904.3642453

[64] Matin Yarmand, Chen Chen, Danilo Gasques, James D Murphy, and Nadir Weibel. 2021. Facilitating remote design thinking workshops in healthcare: the case of contouring in radiation oncology. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–5.

[65] Matin Yarmand, Srishti Palani, and Scott Klemmer. 2021. Adjacent Display of Relevant Discussion Helps Resolve Confusion. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 143 (April 2021), 11 pages. doi:10.1145/3449217

[66] Kangyu Yuan, Hehai Lin, Shilei Cao, Zhenhui Peng, Qingyu Guo, and Xiaojuan Ma. 2023. CriTrainer: An Adaptive Training Tool for Critical Paper Reading. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology* (San Francisco, CA, USA) *(UIST '23)*. Association for Computing Machinery, New York, NY, USA, Article 44, 17 pages. doi:10.1145/3586183.3606816

[67] Tyson Yunkaporta. 2019. *Sand Talk: How Indigenous Thinking Can Save the World*. Text Publishing.

[68] Chengbo Zheng, Yuanhao Zhang, Zeyu Huang, Chuhan Shi, Minrui Xu, and Xiaojuan Ma. 2024. DiscipLink: Unfolding Interdisciplinary Information Seeking Process via Human-AI Co-Exploration. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology* (Pittsburgh, PA, USA) *(UIST '24)*. Association for Computing Machinery, New York, NY, USA, Article 91, 20 pages. doi:10.1145/3654777.3676366

A Survey Questions

Each subsection presents the questions on every page of the survey.

A.1 Set-up

In the next page, please read a short passage from one of Lucy Suchman's influential works in the space of technology and society. This passage is the first two paragraphs of the Introduction section. We would like to gauge your level of comprehension of this write-up.

> Lucy Suchman is an Emeritus Professor at the Lancaster university. Her research intersects feminist science and technology studies, extending her long standing critical engagement with the fields of artificial intelligence and human-computer interaction. Lucy led the Work Practice and Technology group at Xerox Parc for 20 years, and received the SIGCHI Lifetime Research Award in 2010.

A.2 Reading

> This essay takes up the question of cultural practices in the production and use of technical systems, and of what some alternative approaches to our understanding and development of those practices might be. My starting place is recent moves to reframe objectivity from the epistemic stance necessary to achieve a definitive body of knowledge, to a contingent accomplishment of dynamic processes of knowing and acting. I will argue that these reconceptualizations of objectivity are relevant to our thinking about technologies insofar as technologies comprise the objectification of knowledges and practices in new material forms. Of course the story is more complicated than that, as relations of human practice and technical artifact become ever more layered and intertwined. At the same time that the technological project is one of congealing and objectifying human activities, it is increasingly also one of animating and finding subjectivity in technical artifacts. The assimilation of lived experience to technique goes both ways, which only makes the project of re-imagining technological objects the more urgent.
>
> The discussions on which I propose to draw involve, among other things, a shift from a view of objective knowledge as a single, asituated, master perspective that bases its claims to objectivity in the closure of controversy, to multiple, located, partial perspectives that find their objective character through ongoing processes of debate. The premise is that the latter is not only a better route to objectivity, but that it is in actuality the only way in which claims to objectivity are or ever could be grounded, however much the lived work of knowledge production is deleted from traditional scientific discourse. The feminist move in particular reframes the locus of objectivity from an established body of knowledge not produced or owned by anyone, to knowledges in dynamic production, reproduction and transformation, for which we are all responsible.

A.3 Reflection

(1) How likely are you to continue reading the rest of this paper? You have read 1/28 of the paper so far.

- Very Likely
- Likely
- Neutral
- Unlikely
- Very Unlikely

(2) **How was your experience reading this passage?** You could reflect on your experience by considering criteria such as: writing style, concepts and terminologies, length, etc.

A.4 Set-up

In the next page, please engage with the same passage, as well as the alternative explanations provided on the right side. We would like to evaluate the level of helpfulness of these explanations for the comprehension of the original text.

A.5 Alternative Explanations

Original paragraph:

This essay takes up the question of cultural practices in the production and use of technical systems, and of what some alternative approaches to our understanding and development of those practices might be. My starting place is recent moves to reframe objectivity from the epistemic stance necessary to achieve a definitive body of knowledge, to a contingent accomplishment of dynamic processes of knowing and acting. I will argue that these reconceptualizations of objectivity are relevant to our thinking about technologies insofar as technologies comprise the objectification of knowledges and practices in new material forms. Of course the story is more complicated than that, as relations of human practice and technical artifact become ever more layered and intertwined. At the same time that the technological project is one of congealing and objectifying human activities, it is increasingly also one of animating and finding subjectivity in technical artifacts. The assimilation of lived experience to technique goes both ways, which only makes the project of re-imagining technological objects the more urgent.

Alternative Explanation:

Imagine building a vast library where each book not only captures the wisdom of its authors but also lives and breathes with their essence. Initially, our goal is to arrange the library in a way that claims to hold all the answers in neat, objective rows. But we begin to see that true understanding isn't in static shelves but in the dynamic dance between readers and books. This library isn't just a storehouse of knowledge; it transforms human experiences into bound volumes, while these volumes, in turn, influence and enliven the readers. The intertwining of human stories and technical pages becomes more complex, as each interaction breathes new life into the artifacts, making it imperative to continuously re-imagine what this living library should be.

Figure 8: Screenshot of the survey with the first paragraph of the original passage (left) and alternative explanation (right).

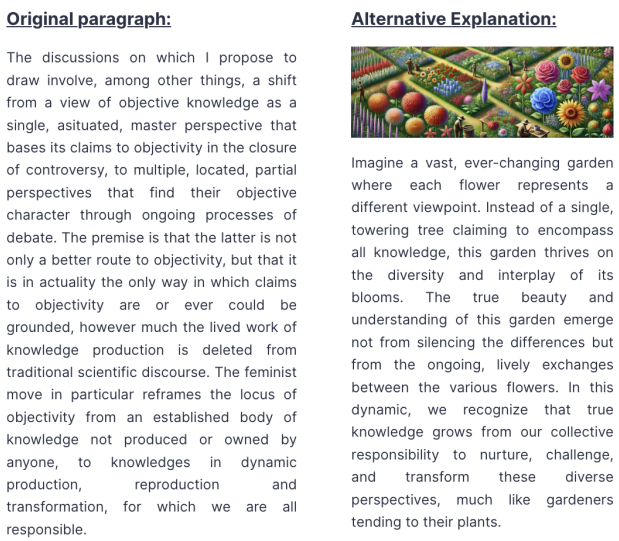


Figure 9: Screenshot of the survey with the second paragraph of the original passage (left) and alternative text (right).

Figure 10: Screenshot of the survey with the first paragraph of the original passage (left) and the adjacent text entry (right).

A.6 Reflection

(1) How likely are you to continue reading the rest of this paper, if you had access to alternative explanations like the one you saw? You have read 1/28 of the paper so far.

- Very Likely
- Likely
- Neutral
- Unlikely
- Very Unlikely

(2) **How was your experience reading the original passage with the alternative explanations?** You can think through: how the text- and image-based explanation helped/hindered your understanding of the original passage, how you related to the explanation, etc.

A.7 Authoring (Figure 10)

Re-explain the original paragraph in your own words. How would you tweak/modify/overhaul the content of the previous alternative explanation in order to help a first-time reader understand the original paragraph?

A.8 Sketching

Design your ideal "Lucy Suchman-reading-helper-system". Imagine you continue reading Lucy Suchman's paper, and the PDF tool (that you are using) can prepare and provide any sort of help (including the alternative explanations provided before) to help with your reading. What would this PDF reader look like? More specifically, how would you want to interact with this text and alternative explanations to better tailor the provided support to your needs?

- To refresh your memory, your previous thoughts on the original passage from Lucy Suchman, as well as with the alternative explanations are provided below.
- Don't worry about creating a professional-looking design! A quick sketch/drawing that illustrates the essential elements of your interface would be sufficient. You can even choose to insert text boxes in place of complex drawing components.
- There are different ways for creating your designs:
  (1) draw on a piece of paper, take a picture, and upload,
  (2) use an online drawing tool (e.g., sketch.io), download your image, and upload, or
  (3) use the provided drawing widget which is convenient, yet limiting in some features such as the lack of selecting and moving objects.

A.9 Demographics

(1) How old are you?
(2) What is your gender?
(3) What is your primary educational institution?
(4) How difficult do you find each of the following categories of the English language in everyday use (Reading & Writing)? [Options: none (I find it super easy), very little, some, average, more than average, much, extreme (I find it super difficult)]
(5) How long have you engaged with HCI research?
(6) What was your background(s) before starting HCI research?

B Metaphor-based Explanations provided by Eight Participants

ID	Metaphor-based Explanation
P3	Imagine a **vast library** where each book not only captures a vast amount of knowledge. We aim to arrange the books in an ordered way that can always give objective knowledge to all our requirements. But we begin to understand that *true objective knowledge is not derived from a single book or a part of a book, but rather a dynamic dance between many books, their authors, and the readers. The library isn't just a storehouse of knowledge but a projection of human experiences, emotions, and explorations on paper.* This intertwining of stories and technical artifacts becomes more complex as each new book is added. The library is a living and breathing being that is continuously evolving as humans participate with it.
P22	You might think of a **library**, as its classic form, set of books structured in static objective shelves. Initially, we thought of true knowledge as such. Though, *true knowledge comes through the dynamic interaction between the books and the readers; where the readers transform their knowledge and their experience into the books, and in turn, the books influence the readers.* This new way of thinking about knowledge, adds more complexity as each interaction adds new dimensions to this library.
P23	Cultural practices shape technology similar to how a chef's background influences their **cooking** style. *Our understanding of objectivity can change from a fixed recipe to a dynamic cooking process.*
P26	Lets think of experiencing a phenomenon with **our senses**. The multiple senses help us gaining the essence of the realities presented while experiencing the phenomenon. *Imagine if you are experiencing the phenomenon only by the sense of sight, your knowledge is limited and hence the idea of objectivity is narrowed as well. Appreciating the multiplicity and dynamic nature of existence and experience of knowledge is key in understanding the world.*
P33	This essay discussed the relationship between practices in production and the use of technical systems. Think about a **giant library** with millions of books. Book authors conceptualize knowledge, construct a body of knowledge, find examples, and reframe knowledge into a book for audiences to read. Audiences may use this knowledge in various scenarios, transforming it in different forms, and applying it to technology development. *As readers explore these books, the knowledge and concepts inspire them to develop new technologies and concepts, which can then be documented in books, further expanding the library. This library dynamically transforms human experiences and practices into new books, and these books, in turn, inspire the readers.* The relationship between the library and the readers is similar to the link between technique and practical experiences, where the development of techniques and living experiences complements each other.
P38	So let's consider the relationships between **rivers and ocean**. The oceans are made up of rivers, and *the direction and speed of each river determines the size of the ocean. The interactive state of river origins and the meeting points of different rivers also influences the state of large rivers.*
P43	Technology is like **cooking**. This essay explores how our cultural practices, as recipes, and technical systems, as utensils are combined to create new dishes (technologies). Before, we thought there was a perfect recipe, but now we know that cooking is a dynamic process. The authors argue that these new ideas are important because *technologies are like dishes that combine knowledge and practices in material forms. The relationship between people and tools is complex, as they influence each other.* Therefore, it is urgent to rethink our recipes and utensils.
P47	Think about objectivity as the way the knowledge is organized, such as the way **libraries** organize books. *The dynamic needs and perspectives of readers (subjectivity) in the library will influence how the organization should be.*

Table 2: Eight of the participants provided metaphoric explanations in the writing task of the survey, when asked to re-explain the original paragraph to a first-time reader. The bolded phrases represent the settings for the metaphors, and the *italicized* parts describe the "dynamic processes of knowing and acting" as explained by Lucy Suchman [57] within the introduced settings. Four of the explanations used *library* (same setting as the provided metaphor), two participants (P23 and P43) selected *cooking*, P26 explained via *humans senses*, and P38 built a metaphor using *rivers and oceans*.