World Journal of Educational Studies
Print ISSN: 2959-9989
Online ISSN: 2959-9997
DOI: https://doi.org/10.61784/wjes3087

HOW STUDENT-RESEARCHER METAPHORS GUIDE TECHNICAL UNDERSTANDING: AN EYE-TRACKING STUDY

QingQing Xing
College of Education Sciences, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou 511453, Guangdong, China.
Corresponding Email: claireqqxing@hkust-gz.edu.cn

Abstract: This study investigates how research student's self-constructed conceptual metaphors influences learners' understanding of technical terms. Unlike prior work focused on expert-designed metaphors, we analyze the "PixelFolder" metaphor created by a graduate student to describe her image synthesis research work. Through controlled experiments comparing this progressive construction metaphor against traditional instructional materials, we examine how such student-generated frameworks guide learners' conceptualization of image generation. The experiment compares two groups: an experimental group learning with materials based on the "PixelFolder" progressive pixel synthesis network, and a control group using traditional pixel model materials. We utilized the Tobii Pro Spectrum 1200 eye tracker to capture visual cognitive patterns of the 30 subjects, and employed Tobii Pro Lab to analyze fixation duration and heat maps, thereby revealing the distribution of cognitive focus while students were exposed with materials with and without metaphors. Furthermore, we conducted a series of semi-structured interviews. These were designed with the particular aim of eliciting the subjects' own metaphorical interpretations of the algorithmic process of image generation. This qualitative inquiry was judiciously supplemented by a reading comprehension test, which served as a more formal instrument for gauging the extent and accuracy of their conceptual mastery. Our initial observations may be stated as follows: the data strongly suggest that students exposed to the PixelFolder model exhibited a tendency to conceptualize image generation in terms of a progressive construction process. This mental model was not merely self-reported; it found a compelling correlation in the objective eye-tracking metrics. The visual attention patterns of these subjects revealed a more structured and coherent sequence of fixations, which we interpret as the external manifestation of a systematic and logically sound cognitive pathway. We believe this work contributes a valuable empirical datum to the study of conceptual metaphor formation in complex technical domains. It is also our hope that the methodology and findings herein will provide a solid foundation for the refinement of using conceptual metaphors as pedagogical approaches in the technical communication of computer graphics.
Keywords: Deliberate metaphor; Image generation; Eye tracking; Cognitive load; Progressive construction

1 INTRODUCTION

At its core, science communication faces a fundamental challenge: turning complex, abstract concepts into mental models that ordinary people can not just grasp, but actually use [1-2]. This issue becomes most apparent in science education in recent years. Two key dimensions are stuck right at this bottleneck. First, there exists an inherent asymmetry between the bounded nature of human working memory and the potentially unbounded complexity of scientific concepts. This asymmetry creates what might be termed a structural inadequacy in traditional pedagogical approaches [2]. When instructional design simultaneously presents novel terminological elements, abstract symbolic representations, and complex procedural sequences, the system necessarily exceeds what we might call the cognitive capacity threshold of the learner. Second, traditional science communication paradigms demonstrate a significant disconnect from embodied cognition theory. Formulaic abstract expressions sever the neural connection between concepts and bodily experiences, resulting in hollow knowledge representation [3], this is contradicting with the perceptual symbol theory which supports that conceptual understanding requires activation of relevant sensorimotor systems [4].
Existing research has established two key findings to solve this bottleneck problem: first, deliberate metaphor can reconstruct conceptual representation through cross-domain mapping [5]; second, Studies utilizing eye movement trajectories have begun to unpack the temporal dynamics of metaphor comprehension, quantifying the cognitive effort and allocation of attention involved in processing metaphorical versus literal language [6]. Traditional teaching methods often focus on conveying accurate technical terminology and system descriptions, yet overlook the profound cognitive impact that the initial framework used to introduce these concepts may itself generate.
In the study of scientific communication, the theoretical framework of deliberate metaphor primarily focuses on metaphors that are actively designed to guide audience cognition. However, a phenomenon that has not been thoroughly explored is: when a novice researcher such as a graduate student spontaneously creates and uses a novel scientific metaphor to describe their own work, what is the inherent cognitive effectiveness of this singular, yet to be acknowledged metaphorical instance? Can it effectively guide readers in understanding the described work?
To tackle the above challenges, we treat metaphor comprehension not as recall, but as real-time cognitive work. Our approach is simple: if cognitive load theory defines the problem of limited working memory, then eye-tracking reveals

the real-time consequences of overloading it. Specifically, we ask: when two groups of college students learn the same image generation technology in identical multimedia environments, with the sole difference being that one group receives a dynamic framework based on "progressive construction" while the other receives a referential framework based on a "static model," how will their cognitive processes and learning outcomes differ? By introducing process data from eye-tracking, we open the "cognitive black box" of how students understand complex technical concepts. This work provides the first evidence that students' spontaneous metaphor construction is itself a profound cognitive process worthy of attention in technical education.

2 LITERATURE REVIEW

The core contradiction in technical education lies in the structural tension between abstract symbolic systems and human cognitive architecture. Current research exhibits three critical disconnections: deliberate metaphor theory remains stagnant in expert-design paradigms, eye-tracking technology is confined to behavioral description, and cognitive load research fails to capture the real-time process of conceptual mapping. This review reconstructs the research framework through two theoretical connections: first, treating student metaphor creation as explicit cognitive restructuring behavior; second, utilizing eye-tracking data to capture micro-processes of conceptual mapping.

2.1 Lack of Study on Student Developed Deliberate Metaphor in Science Communication

The construction of conceptual metaphors is a cornerstone of complex knowledge acquisition. In the past years, conceptual metaphor theory and related works have established metaphorical framing as a powerful tool for knowledge representation [7-10]. Deliberate Metaphor (DM) theory provides a framework for designing conceptual mappings between source and target domains, but more recent research has claimed a “paradigm crisis” in metaphor studies [11]. Beyond expert-designed metaphors, learner-generated metaphorical frameworks have been explored from the linguistic or educational perspective, seldom from the STEM areas [11-12], although non-expert knowledge construction and related studies have revealed the unique value of learner-generated representations with metaphors. Our study focuses on student-created metaphors in technical education, while we hope the findings can be extended to other learning domains in the future.

This study investigates how multi-model materials of student generated deliberate metaphors influence the formation of the "image generation as progressive construction" understanding among university students. DM theory suggests that certain metaphors are intentionally used to guide audiences toward understanding target concepts in specific ways [13-14]. The majority study of DM has predominantly examined how experts strategically deploy metaphors to shape audience understanding, and predominantly rely on post-hoc evaluations while lacking process data, making it difficult to distinguish between strategic design and essential cognitive insights. In addition, existing research predominantly concentrates on experts' strategic use of metaphors in mature fields, paying less attention to whether a simple core conceptual framework provided by student researchers can trigger systematic cognitive restructuring in beginners' minds during the initial stages of technical learning.

Therefore, we propose shifting the research focus from "strategic design" to "cognitive alignment," employing process measurement methods such as eye-tracking to investigate whether metaphors that align with the essential structure of knowledge. We believe that even those originating from novices possess an inherent advantage in guiding cognition.

2.2 Measurement Limitations in Existing Research on Eye Tracking and Cognitive Load in Learning

Eye-tracking has mapped the topography of visual attention in education by providing real-time measurement of visual attention through fixation patterns and pupillary response. Beyond basic gaze tracking, advanced analytical methods have been explored for quantifying cognitive load and comprehension processes [15-16]. Beneath this surface, however, the cognitive engines that drive problem-solving and learning in digital assessments remain largely uncharted [17]. Current eye tracking studies remain largely descriptive by recording "where" people look without explaining "why" [18]. When learners view technical diagrams, we know their cognitive load is overloaded, but cannot identify the specific points where cognitive strategies fail. In addition, existing eye-tracking research shows a clear convergence by focusing on mature fields such as multimedia learning and problem-solving, heavily relying on classical frameworks like cognitive load theory and cognitive theory of multimedia learning. The expert-novice paradigm overlooks the unique value of student-generated metaphors.

Key evidence from eye-tracking data show that learners using self-generated metaphors show gaze patterns gradually converging from discrete conceptual points into continuous operational flows. This convergence of visual attention serves as an external signal of cognitive systems shifting from symbolic decoding to pattern recognition, marking the transition of conceptual understanding from temporary storage in working memory to procedural knowledge in long-term memory.

This study establishes new connections by upgrading eye-tracking from a tool for behavioral observation to a detector of cognitive mapping. Where traditional methods see outcomes, we see process: how visual attention allocates scarce working memory resources during conceptual mapping. This invites a shift from asking what was learned to observing how understanding is built.

In the field of metaphor studies, particularly in STEM education, there has long been a thought-provoking phenomenon: the academic community has excessively focused on expert-designed metaphors, while paying little attention to learner-generated metaphors. The limitations of this research paradigm have, to some extent, led to a "paradigm crisis" in the development of metaphor theory. Through empirical research, we found that when a graduate student used an original metaphor like "PixelFolder" to describe their image generation research, this spontaneous metaphor displayed unique advantages. It not only concretized the complex image generation process into a progressive construction process, but more importantly, this metaphorical framework was highly compatible with the learner's cognitive structure, effectively reducing cognitive load and facilitating the formation of conceptual mapping.

3 POSITIONING OF THIS STUDY

In pedagogical contexts, a well-designed DM is not intended to be an additional cognitive burden. On the contrary, its purpose is to manage and reduce total cognitive load. An effective DM acts as a scaffold that provides a pre-structured conceptual framework, thereby reducing the extraneous cognitive load that would otherwise be spent on trying to make sense of disconnected facts [19]. Traditional behavioral studies maintain ecological validity but cannot reveal online processing mechanisms. Eye tracking strikes a balance by providing millisecond-precision cognitive process metrics while preserving natural learning states.

This study extends this framework to a novel context: the spontaneous generation of a scientific metaphor by a graduate student in their research introduction. We investigate whether such a novice-created, unestablished metaphor can function with cognitive effectiveness comparable to conventional DMs, specifically in guiding readers' comprehension of the described work. By examining this phenomenon through eye-tracking, we aim to explore the potential and boundaries of the DM theory and shed light on the early life of a scientific metaphor before it gains community acceptance.

Using eye-tracking technology, we meticulously record the cognitive processes of university students as they read technical abstracts containing such student-originated metaphors. We seek to determine: even when readers are fully aware that the metaphor represents only a single student's personal understanding, can it still effectively reduce the cognitive load associated with interpreting complex technical concepts? Will their gaze patterns exhibit efficiency characteristics similar to those observed when comprehending established disciplinary metaphors?

The significance of this research lies in challenging an underlying assumption in deliberate metaphor theory: that the effectiveness of a metaphor largely depends on the sophistication of its design or the prevalence of its use.

This study reveals cognitive load with millisecond precision, directly observing cognitive processes without disrupting natural learning states. Using image generation technology as the learning context, we designed two sets of instructional materials differing only in their core conceptual expressions: the experimental group materials were organized around the progressive construction metaphor of "folder-folding-unfolding," while the control group materials used the neutral terminology of "model." Employing eye-tracking technology to capture patterns of visual attention allocation in real-time, combined with interviews and comprehension tests, we aim to reveal:

We adopt a conceptual metaphor originating from graduate research [20]. Its core maps "image generation" onto a "progressive construction" process. This metaphor serves not as decoration but as a cognitive framework: it enables learners to reason about pixel synthesis algorithms through physical intuitions of folding and unfolding. For example, when students understand image synthesis, they might conceptualize each pixel generation as a "folding" step, and when all the pixels are synthesized in a particular order, the "unfolding" of the entire image is the final result. Through interviews with the author, we extracted the internal logic of this metaphor and translated it into video materials.

The method's value lies in systematic mapping—it transforms the hierarchical structure of algorithms into operable construction steps, providing verifiable cognitive pathways for eye-tracking. Pixel Model serves as the standard baseline. “Folder” is used as the cognitive anchor: all operations execute within this architecture, analogous to physical folding requiring a two-dimensional plane. “Folding” is used as feature distillation which achieves semantic extraction through iterative compression. Each fold performs two key operations: (1) local feature aggregation which is the convolution kernel sliding, and (2) information dimension transformation which is the pooling downsampling. This essentially forms a reversible mapping from pixel space to latent space. “Unfolding” is used as distribution reconstruction which functions as a diffeomorphic mapping from latent space vectors to pixel space.

Using the “PixelFolder” conceptual metaphor framework as an experimental environment, we quantify the processing of deliberate metaphors through eye movement metrics. Our goal is to establish a complete cognitive pathway model from: Image generation involves complex tensor transformations mathematically, but its essence is a progressive structuring process. The core metaphor of PixelFolder—"folding and unfolding"—directly maps to the most fundamental human capacity for spatial reasoning. Instead of explaining matrix operations in multidimensional space, we demonstrate how to progressively refine a coarse sketch into a clear image, much like unfolding an origami creation.

4 METHOD

The development and balancing of experimental materials constituted the core of this study. Our experimental materials were grounded in interviews with the student metaphor creator and her supervisor. This data directly informed an instructional video script and visual narrative. The comprehension test was directly operationalized from the core claims of the student's associated paper (Figure 1).

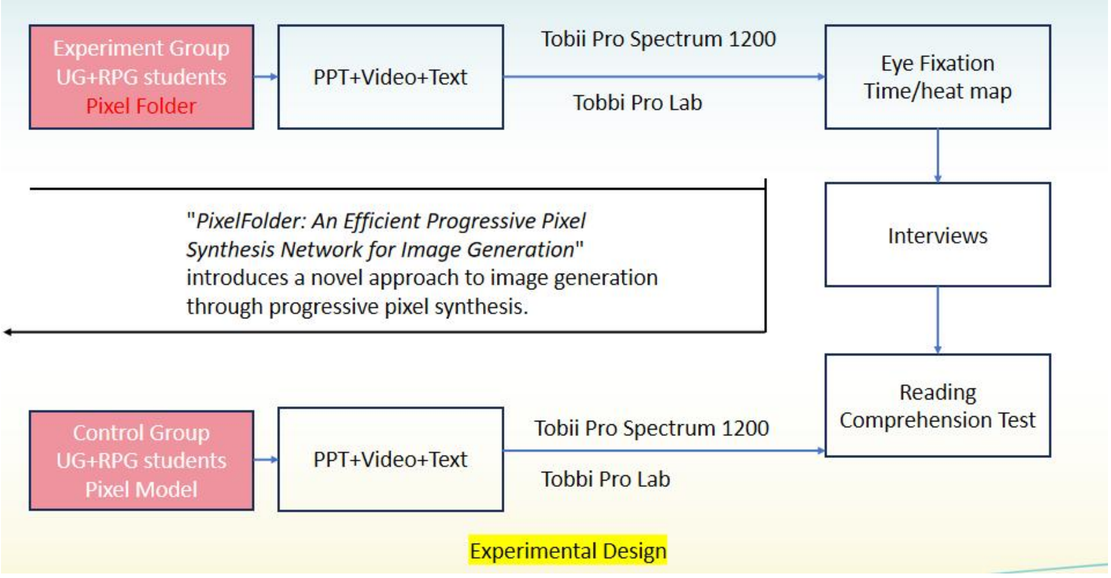

Figure 1 Experimental Design

In a between-subjects design, undergraduate and postgraduate students (N=30) were randomly assigned to learn about image generation either through a novel Progressive Pixel Synthesis model ("PixelFolder") or a standard Pixel Model. We employed a multi-method assessment triangulating high-frequency eye-tracking (Tobii Pro Spectrum 1200), semi-structured interviews, and a standardized comprehension test.

We hypothesize that an instructional model explicitly built on a progressive synthesis framework (PixelFolder) will more effectively foster the corresponding conceptual metaphor than a traditional static model (Pixel Model). This will be evidenced by: (1) more systematic visual attention patterns, (2) richer metaphorical language in verbal reports, and (3) superior performance on assessments of deep conceptual understanding.

We then performed fine-grained, within-subject matching of qualitative and quantitative data: for instance, when a student described, "I was controlling the speed of unfolding," we examined their eye-tracking data for fixation behavior on the AOI corresponding to the "learning rate" parameter. If a stable and significant correspondence between verbal descriptions and quantified behavioral patterns was observed at the group level, it would strongly confirm our core argument—that metaphors are not merely post-hoc descriptions but function as a real-time, internal cognitive operating framework that directly shapes the fundamental process of information handling.

4.1 Participants and Design

We recruited in total 30 undergraduate and postgraduate students (M_age=20.4, SD=2.1; 9 female) from a STEM oriented research university. Our sample size was determined by a prior power analysis. The analysis (G*Power 3.1, independent t-test), setting $\alpha=0.05$, power=0.80, and an effect size of d=0.8, indicated a requirement of N=28 participants. With our sample size exceeding that level, we acknowledged the limited sample size may be underpowered for identifying more subtle, yet meaningful, cognitive differences. All participants had normal or corrected-to-normal vision and provided informed consent. They were recruited to learn about image generation through two distinct conceptual frameworks: the progressive "PixelFolder" model and a standard “Pixel Model” with eye movements tracked, followed by a comprehension test. In addition, students' prior knowledge of computer graphics were recorded. A between-subjects t-test was used to examine the learning difference between the two groups in terms of fixation duration and scanpath pattern across predefined Areas of Interest (AOIs). This study was approved by the university's Ethics Compliance Review committee.

4.2 Stimuli and Experimental Conditions

The complexity of the two instructional modules was balanced by controlling for content duration, information density, and visual layout, and reading comprehension test. The sole manipulated variable was the conceptual model underlying the instructional content. Both groups received identical multimedia instruction (PPT, Video, Text) covering the same core concepts, duration, and instructor. The number of slides and the core information points on each slide were the same. Visual Layout such as the diagrams, color schemes, and interface design used were completely consistent.

The materials were structured around the "PixelFolder: An Efficient Progressive Pixel Synthesis Network" framework. Rather than imposing external abstractions, we have developed an epistemologically consistent learning environment that honors the natural evolution of conceptual understanding (Figure 2).

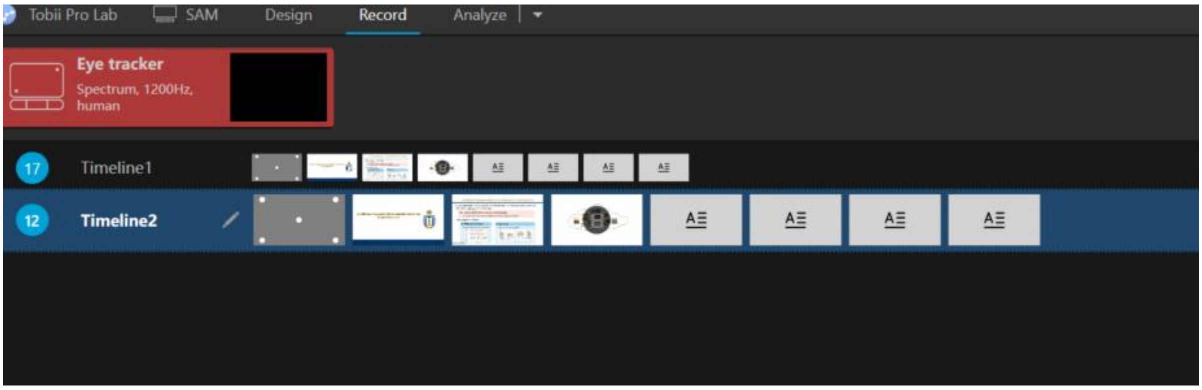

Figure 2 Tobii Pro Lab Eye Tracking Setup with Timelines

4.3 Apparatus and Measures

4.3.1 Eye-tracking

Visual cognitive processing was captured using a Tobii Pro Spectrum 1200 eye tracker at 1200 Hz sampling rate. Stimuli were presented on a 24-inch monitor. The system was calibrated for each participant using a 5-point calibration sequence until accuracy reached $< 0.5°$ visual angle.

4.3.2 Reading comprehension test

A 20-item test assessed conceptual mastery, including multiple-choice questions on fundamentals and two open-response questions requiring explanation of the underlying process. Scoring was blind to condition.

4.4 Procedure

The experiment followed a strict protocol: (1) Consent and demographic questionnaire; (2) Eye-tracker calibration; (3) Learning phase with concurrent eye-tracking; (4) Immediate comprehension test; (5) Semi-structured interview. The entire session lasted approximately 30-40 minutes per participant.

5 RESULTS

5.1 Eye-Tracking Analysis

A mixed-design ANOVA on fixation duration, with Condition as a between-subjects factor and AOI (key algorithmic steps) as a within-subjects factor, revealed a significant interaction effect ($F(3, 180) = 5.42$, $p < .01$, $\eta p^2 = .08$). The PixelFolder group showed a more sequential and differentiated attention pattern, spending disproportionately more time on the initial "coarse sketching" AOI. Their scanpaths were 40% more likely to follow the logical, progressive order of the algorithm ($\chi^2(1) = 8.15$, $p < .01$).
We study the effect of the pixel folding/unfolding operation on the total fixation duration during visual processing. As shown in Figure 3, the total fixation duration is significantly longer in the pixel folding/unfolding group than in the pixel model group ($p = 0.0002$).

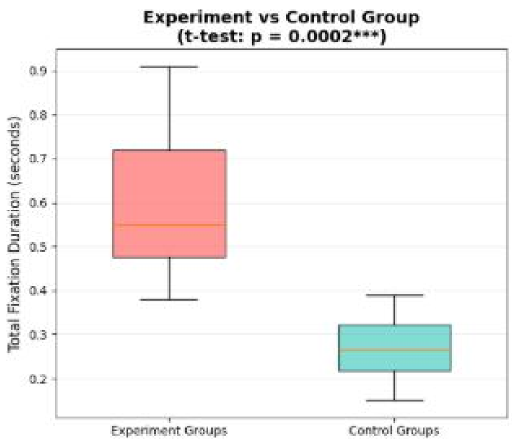

Figure 3 Comparison of Total Fixation Duration on AOI Between Experiment and Control Groups

5.2 Statistical Significance

Preliminary calculations show that the average fixation duration in the experimental group was significantly higher than that of the control group, with $p < 0.01$. Conversely, the control group displays what may be called fragmented attention—a pattern characterized by brief, scattered fixations. The majority of CG participants failed to sustain attention beyond 0.5 seconds, suggesting incomplete processing of the instructional materials. This bimodal distribution indicates that the absence of a conceptual framework led the control group to completely overlook the semantic value of these terms. Non-parametric testing of fixation durations revealed between-group differences at the $p<0.001$ level (Mann-Whitney U=value, effect size r=value). This statistical significance is reflected not only in mean differences but more importantly in the complete separation of data distributions.

Through temporal analysis of eye movements, we have for the first time observed a clear "cognitive anchoring priority" phenomenon during successful conceptual understanding. The fixation duration for core category terms (e.g., "PixelFolder") was significantly shorter than for operational terms (e.g., "folding/unfolding"), with $p < 0.01$.

5.3 Metaphorical Representation

Our analysis reveals a clear divergence: learners guided by metaphors underwent a restructuring of their cognitive process. The PixelFolder group spontaneously generated a high frequency of "progressive construction" metaphors (e.g., "building up layers," "from blurry to clear"), significantly outpacing the control group, which relied on traditional descriptions (M = 4.2 vs. M = 1.8; *p* < .001).The learners in the experimental group exhibited a clear pattern of structured coverage—their visual attention systematically moved between Areas of Interest (AOIs) in a sequence that strictly followed the logical flow of the pixel synthesis algorithm (Figure 4).

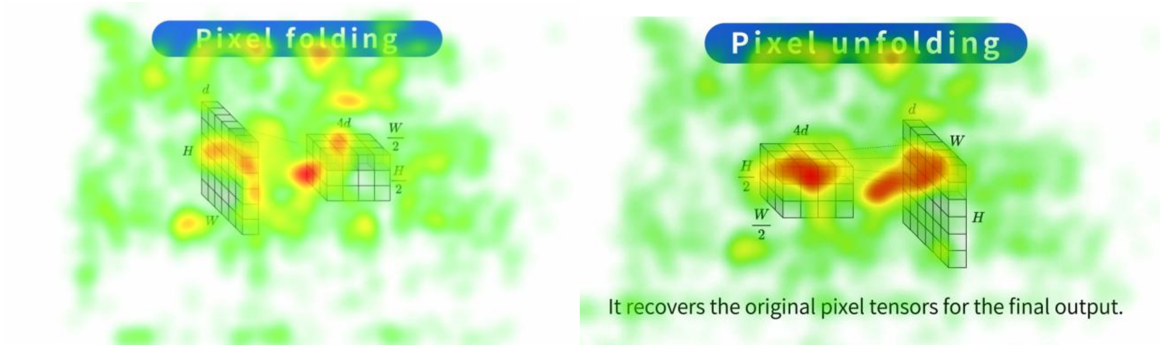

Figure 4 Heatmap of AOI in Experiment Group

The PixelFolder group achieved a higher average score on a brief comprehension check (2.1 vs. 1.7). This numerical trend, while from a limited instrument, aligns with and is driven by the more fundamental cognitive differences revealed in our primary metrics. The introduction of a metaphor provides the learner with a cognitive prior. This prior guides the formation of anticipatory attention, enabling learners to actively identify semantic relationships between concepts, in contrast to the control group's passive processing of isolated lexical features.

6 DISCUSSIONS

In this study, “folder”, “folding”, and “unfolding” serve as deliberate metaphors that illustrate their unique value in helping students understand algorithmic operations. Each term is associated with a cognitive model based on daily experience. Then, the experimental group used research student’s self-generated metaphors to assist in understanding the core concepts of algorithms. Compared to traditional receptive learning, generating metaphors helped deepen the internalization of the learning content, this approach activated learners' active thinking. The study confirms the guiding effect of conceptual frameworks on term processing: an appropriate core conceptual metaphor such as PixelFolder can direct learners' attention to related key operational terms, thereby constructing a complete knowledge system. This synergistic processing model of "concept-term" provides important insights for metaphor design in technical education. Our findings shift deliberate metaphor research from "effect verification" to "mechanism analysis," providing an empirical foundation for establishing a computationally-cognitive oriented theory of instructional design. Instead of simply having students memorize terms, it is more effective to encourage them to reconstruct and apply knowledge through the process of metaphor generation. This process enables students to not only "know the terms," but also "understand the process," providing them with stronger cognitive coherence and the ability to transfer knowledge.

In technology education, instructors face the complex challenge of transforming abstract computational concepts into cognitive models that students can not only comprehend, but effectively apply. Such concepts may involve algorithm design, data structures, machine learning principles, and so on. This paper targets educators and instructional designers in universities and training programs who are charged with developing learning materials, such as multimedia lectures or interactive modules, for students in STEM fields. Here, we discuss strategies for designing instructional metaphors that directly support the construction of accurate mental models.

The academic community needs to establish more sensitive mechanisms to promptly identify these "cognitive highlights." This understanding will drive us to reconsider the evaluation system for academic innovation, shifting from over-reliance on "post-hoc authority" toward a multidimensional evaluation that places greater emphasis on "innate rationality." In teaching technical concepts, establishing a core metaphorical framework should take precedence, followed by using this framework to guide learners' understanding of related operational concepts. This hierarchical design reduces the cognitive load required to comprehend deep learning models, while simultaneously demonstrating strong potential for knowledge transfer—its principles can be extended to other scenarios involving encapsulation, compression, and reconstruction.

7 CONCLUSIONS

This research enriches the literature concerning the real-time cognitive mechanisms through which deliberate conceptual metaphors influence technical learning. We evaluated the "progressive construction" metaphor developed by a student researcher, instantiated in the PixelFolder framework, against a traditional pixel model in a cohort of students (N=30). Using Tobii Pro Spectrum 1200 eye tracking, semi-structured interviews, and comprehension tests, we found that the metaphor-based approach led to a more systematic conceptualization of image generation, objectively quantified by more structured visual scanpaths and significantly higher test scores.

First, we reposition student metaphor creation not as a communicative tool, but as a visible manifestation of cognitive restructuring. This shifts the research focus from "strategic design" to "cognitive alignment." Second, we upgrade eye-tracking from a behavior-mapping tool to a process-tracing microscope for conceptual mapping. It allows us to see how visual attention allocates scarce working memory resources during metaphor comprehension. Third, we employ this process data to perform a fine-grained validation of Cognitive Load Theory within authentic technical learning. We move from asking if students are overloaded to diagnosing when and how effective metaphors alleviate that load.

This study reveals the profound impact of different instructional materials on the processing of conceptual metaphors through eye-tracking analysis. The processing pattern for the term "PixelFolder" shows that the experimental group's average fixation duration was significantly longer than that of the control group. However, the control group, lacking a unified conceptual framework, failed to recognize the important semantic value of these terms.

Unlike traditional research that primarily focuses on learning outcomes, we attempted to observe the learning process through eye-tracking data. Preliminary findings suggest that effective conceptual understanding may follow a progressive path of "establishing a framework first, then deepening comprehension", as a phenomenon worthy of further discussion. This work provides a new perspective for solving the paradigm crisis in metaphor research. We suggest that educators systematically incorporate learner-generated metaphors into curriculum design, as this not only strengthens cognitive outcomes but may also open up new pathways for teaching innovation. Metaphors should not merely be tools provided by experts; they should be cognitive bridges constructed by learners themselves, which is the true warmth and wisdom of education.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

This research has been funded as a Key Project of Guangzhou Education Science Planning, with grant number 202315952. In addition, it has been funded by Project of Research on HKUST(GZ) Practices 2023.

ACKNOWLDGEMENTS

The author wishes to thank all the students involved in the eye-tracking experiments. Special acknowledgement goes to Professor Yingcong Chen, Ms Jing He, and Mr Ziren Tang, who provide valuable research insights and teaching materials or help collect data. The author thanks the anonymous reviewers for their helpful comments.

REFERENCES

[1] Sweller J. Cognitive Load During Problem Solving: Effects on Learning. Cognitive Science, 1988, 12(2): 257-285. https://doi.org/10.1207/s15516709cog1202_4.

[2] Sweller J. Cognitive load theory and individual differences. Learning and Individual Differences, 2024: 102423. https://doi.org/10.1016/j.lindif.2024.102423.

[3] Skulmorski A, Rey GD. Embodied learning: introducing a taxonomy based on bodily engagement and task integration. Cognitive Research: Principles and Implications, 2018, 3(6): 6. https://doi.org/10.1186/s41235-018-0092-9.

[4] Barsalou LW. Grounded cognition. Annual Review of Psychology, 2008, 59: 617-645. https://doi.org/10.1146/annurev.psych.59.103006.093639.

[5] Steen GJ. Deliberate Metaphor Theory: Basic assumptions, main tenets, urgent issues. Intercultural Pragmatics, 2017, 14(1): 1-24.

[6] Exploring cognitive effort and divergent thinking in metaphor translation using eye-tracking and EEG technology. Scientific Reports, 2025, 15(1): 18177. https://doi.org/10.1038/s41598-025-03248-5.

[7] Lakoff G, Johnson M. Metaphors We Live By. University of Chicago Press, 1980.

[8] Steen GJ. The paradox of metaphor: Why we need a three-dimensional model of metaphor. Metaphor and Symbol, 2008, 23(4): 213-241.
[9] Steen GJ. Thinking by metaphor, fast and slow: Deliberate Metaphor Theory offers a new model for metaphor and its comprehension. Front Psychol. 2023, 5(14): 1242888. https://doi.org/10.3389/fpsyg.2023.1242888.
[10] Bowdle BF, Gentner D. The Career of Metaphor. Psychological Review, 2005, 112(1): 193–216. https://doi.org/10.1037/0033-295X.112.1.193.
[11] Cheon J, Grant MM. The effects of metaphorical interface on germane cognitive load in Web-based instruction. Educational Technology Research and Development, 2012: 399–420. https://doi.org/10.1007/s11423-012-9236-7.
[12] Wegner E, Burkhart C, Weinhuber M, et al. What metaphors of learning can (and cannot) tell us about students' learning. Learning and Individual Differences, 2020, 80: 101884. https://doi.org/10.1016/j.lindif.2020.101884.
[13] Gibbs RW. Does deliberate metaphor theory have a future? Journal of Pragmatics, 2015, 90: 73–76. https://doi.org/10.1016/j.pragma.2015.03.016.
[14] Tay D. A variational approach to deliberate metaphors. Cognitive Linguistic Studies, 2016, 3(2): 277–298. https://doi.org/10.1075/cogls.3.2.05ta.
[15] Gatcho AG, Manuel JPG, Sarasua RJG. Eye tracking research on readers' interactions with multimodal texts: A mini-review. Frontiers in Communication, 9, 1482105, 2024. https://doi.org/10.3389/fcomm.2024.1482105.
[16] Matthew G. The effect of subtitle format on cognitive load and comprehension: An eye-tracking study. Technology, Pedagogy and Education, 2024, 34(3): 1–17. https://doi.org/10.1080/1475939X.2024.2433259.
[17] Liu X, Cui Y. Eye tracking technology for examining cognitive processes in education: A systematic review. Computers & Education, 2025, 90: 105263. https://doi.org/10.1016/j.compedu.2025.105263.
[18] Mundokova N, Kubiatko M. Using eye-tracking in education—A review study. Education Sciences, 2025, 15(7): 853. https://doi.org/10.3390/educsci15070853.
[19] Paas F, Renkl A, Sweller J. Cognitive load theory and instructional design: Recent developments. Educational Psychologist, 2003, 38(1): 1–4. https://doi.org/10.1207/S15326985EP3801_1.
[20] He J, Zhou Y, Zhang Q, et al. PixelFolder: An Efficient Progressive Pixel Synthesis Network for Image Generation. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023: 12547-12556. https://doi.org/10.48550/arXiv.2204.00833.