Metaphors: Optional but Preferred for Conveying Inferential Meaning

Anastasia Smirnova (smirnov@sfsu.edu)[1] & Thomas Raz Parker[2]
[1]Department of English Language and Literature, San Francisco State University
[2]Department of Linguistics, University of New Mexico

Abstract

The status and function of metaphors in language and cognition have been a subject of a long-standing debate. We address this question by introducing *inferential affordance*, a theoretical construct that distinguishes linguistic form from the inferential meaning licensed by that form. Metaphors such as *black holes are messy eaters* convey inferences about abstract phenomena. Is metaphor necessary to express such inferences? The results of two studies show that inferential affordances often survive metaphor removal. At the same time, metaphors emerge as a preferred representational strategy. These findings challenge strong versions of Conceptual Metaphor Theory that view metaphors as vehicles for conveying inferential meaning (Lakoff & Johnson, 1980). They are not fully explained by views that treat metaphors as optional rhetorical devices (Pinker, 2007). Instead, our findings align closely with the view of metaphors as analogies (Holyoak, 2019), according to which metaphors are optional but selectively recruited to support inferential understanding.

Keywords: metaphors; inferential affordance; analogy; representational choice

Introduction

Black holes are messy eaters, a metaphor by Cambridge astronomer Mathew Bothwell, characterizes the behavior of complex cosmic objects and their tendency to only partially consume the surrounding matter.[1] Such metaphors illustrate how figurative language can support understanding by enabling inferences about abstract, unobservable phenomena. Understanding of the metaphor is based on establishing the relevant elements and relationships between them in the source domain (food consumption) and transferring these elements and relationships to an abstract domain (cosmic objects).

In this study we ask how metaphors support understanding of relational properties of objects in physical and social environments and evaluate this question with respect to the current debate on the nature of metaphors in cognitive science. We operationalize the ability of metaphors to support understanding in terms of *inferential affordance* – inferences about relations that are licensed by linguistic representation, without expectation that these inferences are explicitly formulated. Our notion of inferential affordance is based on Gibson's (1979) concept of affordance: objects in the physical environments – apple, chair – are perceived by an agent from the perspective of the actions that can be performed on them – eating and sitting, respectively. We differ from Gibson and ecological psychologists in that we apply affordance to the mental process – inference – in line with more recent research on mental affordances (cf. Jorba, 2019; McClelland, 2020).

Inferential affordances of metaphors most prominently come to light in the context of education. According to the analogy approach to metaphor, a position that we adopt here, understanding of metaphors is based on the same mechanisms that underlie analogical reasoning: establishing relational similarities between entities in different conceptual domains (Gentner, 1983; Gentner & Clement, 1988; Holyoak, 2019). To illustrate how inferential affordances arise from metaphor, consider the earlier example. The source domain supplies the eating relation that holds between the eater and the food. In the target domain, the relevant relation is absorption, and it holds between black holes and the surrounding cosmic matter. These relations can be represented in predicate-argument format: DEVOUR (EATER, FOOD) and ABSORB (BLACK HOLE, COSMIC MATTER). Once similarities in the relations DEVOUR ~ ABSORB are established, the reader can identify systematic correspondences between the two domains (EATER corresponds to BLACK HOLE, FOOD corresponds to COSMIC MATTER). The adjective *messy* further implies that similarly to crumbs, spills, and food leftovers around untidy eaters, black holes are surrounded by partially consumed cosmic matter. These are inferential affordances that arise from the metaphor. Collectively, they help the reader to understand properties of cosmic objects.

Teaching difficult concepts by employing analogy and metaphors is a promising direction for education research informed by cognitive science (Gray & Holyoak, 2021). Metaphors help learners understand abstract concepts (Duit, 1991) and can facilitate recall (Harris et al., 1999; see Alfieri et al., 2013; Low, 2008 for overview). However, this approach also raises concerns about cognitive cost. Analogical comparison is cognitively demanding, and there is a possibility that metaphors in novel educational contexts may tap into additional cognitive resources beyond what is normally required by a learning task, thus increasing

[1]https://science-education-research.com/reference/science-comparisons/black-holes-swallow-light/

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

cognitive load (Gray & Holyoak, 2021). These conflicting considerations – metaphors support understanding yet are cognitively costly – create a natural setting for examining which representational choices are preserved when comprehension is a priority. We study how these conflicting considerations play out in the context of text simplification. During simplification, complexity of a text is reduced to facilitate ease of understanding, yet the meaning is preserved.

The question we address is the following: Are metaphors necessary for conveying inferential affordances? We operationalize this question by analyzing representational choices made by educators (metaphorical vs. non-metaphorical form) and examine how these representational choices relate to inferential affordances, i.e. whether affordances survive when metaphors are removed.

Theories of Metaphors

The nature of metaphors and their status in language and cognition have been a matter of a long-standing debate. The proponents of Conceptual Metaphor Theory (CMT) argue that thinking is for the most part metaphorical (Lakoff & Johnson, 1980). Metaphorical foundations of thought emerge from our interaction with the physical world and are based on perceptual experience (embodied cognition). These experiences create foundations for conceptual structure. Concepts and relations among them are represented as conceptual mappings, such as UNDERSTANDING IS SEEING CLEARLY. Experimental studies within the CMT framework suggest that such conceptual mappings are activated during language processing and support understanding of abstract concepts and figurative language (Gibbs & O'Brien 1990; Gibbs et al., 1997; Nayak & Gibbs, 1990). For example, understanding of idioms – *spill the beans* means 'reveal a secret' – is supported by conceptual metaphors MIND IS A CONTAINER and IDEAS ARE OBJECTS.

The CMT framework makes specific predictions about metaphors in contexts that prioritize conceptual understanding but also aim to reduce processing demands, such as simplification. According to this approach, metaphors that instantiate core conceptual relations will be preserved even if they entail an increase in cognitive load. Without them we risk the loss of understanding. Support for this position comes from an empirical study by Wolska & Clausen (2017). The authors study metaphors in educational texts simplified for upper elementary curriculum. By directly comparing the distribution of metaphors in original informational texts and in simplified texts they observed that the majority of metaphors (69%) survive the simplification process. Even though the authors do not directly relate their work to the CMT framework, their findings can be viewed as indirect support for this position. They suggest that metaphorical content performs an essential function and merits preservation despite the potential increase in cognitive load.

A broad alternative to CMT rejects the claim that reasoning is metaphorical in nature and views metaphors as linguistic and pragmatic devises that are external to inferential processes (Glucksberg et al., 1993; Murphy, 1996; Pinker, 2007; Sperber & Wilson, 2008). A clear articulation of this position is Pinker (2007: 250): "Thinking cannot trade in metaphors directly. It must trade in a more basic currency that captures the abstract concepts shared by the metaphor and its topic." In linguistics, Relevance Theory (Sperber & Wilson, 2008) provides a similar view. It shares with Pinker the position that metaphors do not have any special status in cognition and argues that their distribution in language is governed by general principles of language use. Metaphors are linguistic packaging that can express inferential meaning, derived pragmatically; they are rhetorical devices and their presence is optional. The loss of metaphor would not impair understanding because the same content can be expressed literally. This approach predicts that metaphors can be sacrificed in the context of text simplification without the loss of inferential affordances, if the relevant propositional content is conveyed by alternative linguistic means.

In-between the two extreme positions, lies a theoretical perspective that views metaphors as conceptually relevant but optional (Gentner, 1983; Holyoak, 2019). The proponents of this approach would agree with Pinker (2007) that the foundations of thinking are not metaphorical. However, unlike Pinker (2007), they argue that metaphors can do real cognitive work. Under this view, metaphors utilize analogical reasoning. Analogical reasoning operates on structured domains: metaphors activate relational similarities between the source and the target domains, and the relevant relations from the source are transferred to the target (Gentner & Clement, 1988; Wolff & Gentner, 2011). While metaphors are not imposed by conceptual structure, they can be used to facilitate understanding and "make it easier for the reader to grasp what is intended" (Holyoak, 2019: 59). From this perspective, metaphors are not obligatory, but they can be used strategically to support understanding. In the context of text simplification, this position predicts selective preservation.

Text Simplification as Empirical Testing Ground

Text simplification in an educational context aims to reduce structural and lexical complexity of the text while preserving its meaning (Siddharthan, 2014). This process creates pressure on representational choices. Difficult words are replaced by easy-to-understand synonyms (*demonstrated* → *showed*; *sphere* → *round object*) and complex clauses are split into shorter, syntactically simpler sentences (*The skeleton discovered in Utah provides new clues* → *The skeleton was discovered in Utah. It provides new clues*). In this study we focus on what happens to metaphors that carry inferential affordances when text is simplified. Since simplification aims to preserve meaning, a good simplification should convey inferential meaning that captures the relational and causal information present in the original text. If metaphors that carry inferential affordances are mostly preserved, this would provide support for CMT. If

metaphors are dropped but their inferential affordances are preserved, this would support the view of metaphors as rhetorical devices. If metaphors are selectively preserved, especially in contexts that support understanding of relational properties, this would support the analogical view of metaphors as optional but useful representational tools.

Human Educators and LLMs as Two Contrast Cases

Are metaphors necessary to convey inferential affordances or are they merely preferred representational devices? The analysis of representational choices that human educators make can help us identify which representations are preferred. However, human representational choices alone do not suffice to distinguish preference from necessity. In order to address this question, we need to establish a contrast with the system that can perform the same task (text simplification), maintain the same goal (meaning preservation), but might favor different representational choices. To address this consideration, we introduce LLMs as a contrast case. The use of non-human systems as contrast cases is not unprecedented in cognitive science. Research in comparative cognition uses non-human species to isolate cognitive capacity such as problem solving from linguistic system (Thorndike, 1911; see Gómez, 2022 for an overview). We extend this methodological logic to LLMs.

Previous work has shown that LLMs are capable of identifying and interpreting metaphors (Hicke & Kristensen-McLachlan, 2024; Ichien et al., 2024). Thus, LLMs can convey metaphorical meaning, but they might not necessarily treat metaphors as default vehicles for inferential meaning. Representational choices that LLMs make are the ones that make sense in the context of the task (McCoy et al., 2024). These choices might be orthogonal to human communicative or pedagogical norms, and might simply reflect performance driven by constrains imposed by the task. If inferential affordances survive in LLM-simplified texts while metaphors are lost, this would suggest that linguistic metaphor is not necessary to convey inferential meaning. If on the other hand humans retain inferential affordance as well as the metaphor in the same case, this would indicate a clear preference and a stronger association between metaphor and inferential meaning. This contrast allows us to examine whether metaphors and inferential affordances are equally strongly associated in two different systems.

Study 1: Metaphors and Inferential Affordances in Human-Simplified Text

In this study, we address the question of whether metaphors are essential for conveying inferential affordances. We focus on two questions: (i) What happens to metaphors when informational texts are simplified by humans? (ii) What happens to inferential affordances if metaphors are removed?

Data

We study the distribution of metaphors and their inferential affordances in educational texts from Newsela, a major provider of K–12 curriculum (Xu et al., 2015). The Newsela corpus is a parallel corpus of original and simplified texts adapted for different grade and proficiency levels by professional educators. The corpus is considered a gold standard in text simplification literature, and is used as baseline in studies on text complexity. Our analysis focuses on the comparison between original texts and texts simplified for grade 4, since these are the largest subsets of data within Newsela. These two levels were previously analyzed by Wolska & Clausen (2017) in their study on metaphors, which establishes the relevant comparison basis for our work.

Design & Procedure

To compare the distribution of metaphors and their inferential affordances in original and simplified texts, we employ within-item comparative design. For each metaphor with inferential affordances in original texts we analyze whether (i) the metaphor is preserved or lost in the simplified text, and (ii) whether its inferential affordance is preserved or removed. Metaphor preservation and affordance preservation function as dependent variables.

The procedure and methodological choices were driven by specific theoretical considerations and operationalizations of theoretical constructs. First, we focused only on inferential metaphors, ignoring metaphors that have affective or conventional meaning. We operationalized the construct of inferential affordance as causal, predictive, or relational inferences that go beyond what is asserted. For example, in the following excerpt, *top predator* is identified as an inferential metaphor: *The removal of **top predators** has been called "humankind's most pervasive influence on nature."* The inferential affordances it generates are: (i) ecosystems have hierarchical structure; (ii) these systems are directional: removing top predators affects animals lower in the system. On the other hand, the metaphorical expression *going after* in *Going after the more valuable predators first* allows only the surface understanding of pursuit, and does not have any of the inferential affordance associated with *top predators*. It has descriptive and possibly idiomatic function, but it does not give rise to inferences about ecological systems. Second, because inferential affordances might present a challenge for a naïve interpreter and currently cannot be automated, we prioritized the depth of the analyzes rather than the breadth of coverage. This consideration motivates the size of data we analyzed. Finally, since the inferential affordances can be distributed across multiple sentences, the unit of analysis in our data were passages, not individual sentences.

Passage Alignment In order to identify meaningful passages within original texts and the corresponding passages in simplified texts, we implemented the Jiang et al. (2020) Neural Conditional Random Field (CRF) algorithm for passage alignment. The algorithm was specifically designed to align original and simplified passages of the same text and

was trained on the Newsela corpus. A passage can consist of an individual sentence but it also can be longer. Passage alignment is based on the semantic similarity scores (derived from BERT embeddings) and on positional similarity in the text. For detailed description of implementation steps, see Smirnova et al. (2025b). We implemented the alignment algorithm on a randomly selected subset of Newsela texts and their simplified equivalents. We examined a selected subset of the data. Misaligned passages were removed. We also removed cases when alignment consisted of article headers or section titles only. From the resulting set, we randomly selected one hundred aligned passages for the study of metaphor and used them in the subsequent annotation step.

Annotation Procedure The annotation step consisted of three tasks: identifying inferential metaphors in the original text, establishing whether the metaphor was dropped or preserved in the corresponding simplified text, and establishing whether the inferential affordance of the metaphor was preserved or lost. We used the selectional criteria outlined above for identifying metaphors in text. When annotating simplified texts, the annotators checked if the corresponding metaphor was present. Metaphor preservation is a categorical variable; annotators recorded 1 if the metaphor was preserved and 0 if it was dropped. Inferential affordance of the metaphor is also a categorical variable, coded 1 for preservation and 0 for loss.

Annotators, Inter-rater Agreement, and Gold Our annotation task involved understanding of metaphors and their inferential affordances. Because this task requires theoretical expertise, the two authors served as annotators. We developed an annotation guide and independently annotated one hundred aligned passages. We then met to discuss the annotations and resolve disagreement.

We computed percent agreement on the metaphor selection task and inter-rater reliability, as well as percent agreement and inter-rater reliability for affordance preservation.

In the initial metaphor selection task, inter-rater agreement was low (43% overlap; Cohen's κ=-0.378). Given the inherently interpretive and subjective nature of this task, this is expected; this stage concerned the creation of an expert-adjudicated dataset that served as the basis for all subsequent analysis, not evaluation of independent detection, nor was it intended to make claims about metaphor prevalence in the text. This level of agreement is acceptable for this stage. Future versions of this protocol will take advantage of larger rater pools, and operationalization of selection criteria for relevant metaphors based on the findings of this project. After discussion and adjudication, we ended up with 74 instances of metaphors in original texts.

For the task of identifying metaphor and inferential affordance preservation – providing the data for analysis – inter-rater agreement was high (97.30% overlap; Cohen's κ=0.945), indicating strong reproducibility.

Results

Tables 1 and 2 summarize the pattern of metaphor and affordance preservation in human-simplified texts. As shown in Table 1, metaphors are more often removed rather than preserved during simplification. At the same time, inferential affordances survive at a greater rate than metaphors.

Table 1: Metaphor and inferential affordance preservation in human-simplified texts.

	Preserved	Dropped
Metaphor	41% (30)	59% (44)
Inferential affordance	47% (35)	53% (39)

Table 2 further shows that inferential affordances can survive metaphor removal: in 8 cases affordances were preserved despite the loss of the metaphor. On the other hand, cases in which metaphors were preserved but inferential affordances were lost are rather rare.

Table 2: Contingency of metaphor and inferential affordance preservation in human-simplified texts.

	Affordance preserved	Affordance lost
Metaphor preserved	27	3
Metaphor lost	8	38

These results suggest that inferential affordance can be conveyed without metaphor, even though metaphor preservation is strongly associated with affordance preservation in human-simplified texts.

Interim Discussion

Our findings differ from the results reported by Wolska & Clausen (2017), who argue that metaphors are mostly preserved in text simplification. One possible explanation is the methodological difference: Wolska & Clausen (2017) followed Steen's (2010) procedure for metaphor identification and analyzed all types of metaphorical language, including affective and conventionalized metaphors. In our paper, we focused on inferential metaphors only.

Our results have theoretical implications. Under a strong CMT these results present a puzzle. The theory predicts that the loss of metaphor would result in a loss of the corresponding inferential structure. However, our data challenges this prediction by showing that metaphors and inferential affordances are preserved at different rates. This suggests that a dissociation between metaphor and inferential affordance is possible. Table 3 illustrates three cases of inferential affordance preservation despite metaphor removal.

While Study 1 reveals the representational choices that educators make in the context that prioritizes content preservation, clarity and accessibility, it does not determine whether these choices reflect preference or necessity. Study

2 addresses this question by using LLMs as a contrastive case.

Table 3: Affordance preservation with metaphor removal.

Original text	Human-simplified
The number of little brown bats has **plummeted** about 90 percent.	Nine out of every 10 little brown bats have died.
Making their raft more **porous**	created big spaces between each connected ant
self-healing buildings and bridges	buildings and bridges that can fix themselves

Study 2: Dissociation of Metaphors and Inferential Affordances in Humans and LLMs

In this study we compare representational choices by humans and LLMs. The LLM performs the same task – text simplification – and shares the same goal: to convey the content while making it more accessible to the upper elementary reader. While LLMs are able to identify and understand metaphors, they might differ from humans in how strongly metaphors are associated with inferential affordances.

Data

Original (non-simplified) texts from the Newsela corpus that were used in Study 1 were submitted as input to the OpenAI API with instruction to simplify. We used OpenAI API with the following settings: model = GPT-4o, temperature = 1. The prompt specified the target complexity level of the output text (4th grade), the target length (the average length of texts for 4th grade in the Newsela corpus), and the targeted genre. The following prompt was used: "In approximately 680 words, simplify the text below for a fourth grade reading level written in the newspaper genre". We used a zero-shot strategy, so no examples of simplified texts were submitted as part of the prompt. Since text complexity measures used by Newsela, Lexile score, is based on the average sentence and word length, Newsela editors do not explicitly target metaphors. In parallel, we did not include any specific instructions in the prompt regarding metaphorical content. Prior analyses (Smirnova et al., 2025a) have demonstrated that GPT-4o generated texts align with human-simplified texts on the standard measures of text complexity and preserve the factual content of the original texts. We conclude that the two types of texts are sufficiently aligned for the comparison performed here.

Design & Procedure

Passage Alignment The passage alignment algorithm from Study 1 was used to align original and LLM-simplified texts. The three sources of data were merged to create aligned triplets of original, human-simplified and LLM-simplified texts. These one-hundred aligned triplets served as the basis of our analysis.

Annotation Procedure We followed the same annotation procedure as in Study 1, which allows for a direct comparison of metaphors and inferential affordances in human- and LLM-simplified texts. Both authors served as annotators. We evaluated whether inferential metaphors from original texts were preserved in LLM-simplified texts and whether their inferential affordances were preserved.

Annotators, Inter-rater Agreement, and Gold Inter-rater agreement for annotation for metaphor and inferential affordance preservation for the LLM texts was high (90.54% overlap; $\kappa = 0.81$), indicating strong reproducibility.

Results

Tables 4 shows that metaphors are retained by humans and LLMs in comparable rates.

Table 4: Metaphor and inferential affordance preservation in human- and LLM-simplified texts.

	Metaphor retention	Affordance preservation
Human	41% (30)	47% (35)
LLM	42% (31)	54% (40)

Table 5 presents the contingency between metaphor and inferential affordance preservation in LLM-simplified texts, which allows for a direct comparison with the corresponding human contingency in Study 1.

Table 5: Contingency of metaphor and inferential affordance preservation in LLM-simplified texts.

	Affordance preserved	Affordance lost
Metaphor preserved	26	5
Metaphor lost	14	29

We compared instances where inferential affordances were retained in both the LLM and human sentences using McNemar's test. McNemar's test showed no statistically significant asymmetry in preservation ($\chi^2 = 9.0$, $p = .405$). This suggests that instances where inferential affordance was lost in either the human or LLM-simplified texts occurred at approximately the same rate. Importantly, these results indicate that there is no systematic difference in overall affordance preservation rates between humans and LLMs. This motivates a closer look at representational strategies.

To examine representational strategies, a subset of instances where inferential affordance was retained in both the human and LLM-simplified texts was inspected for

metaphor preservation. Humans retained metaphors in five cases where the LLM removed them, while the LLM retained metaphors in two cases where humans removed them. Table 6 shows two examples of this asymmetry.

Table 6: Examples of metaphor retention asymmetry between human- and LLM-simplified texts.

Original text	Human-simplified	LLM-simplified
What other characteristics might it **drive**	What other changes might it **drive**?	Does it also change how they fly or do other things?
The number of little brown bats has **plummeted** about 90 percent	Nine out of every 10 little brown bats have died	This disease has made their numbers **fall** by about 90%

There is a descriptive tendency for humans to rely on retaining metaphors more often than LLMs. Even though this asymmetry is not statistically reliable (McNemar's test, $\chi^2 = 2.0$, $p = .453$), it suggests that humans might be more likely than LLMs to retain the metaphorical expression when inferential content is preserved.

Table 7 provides a more fine-grained perspective on the differences between humans and LLMs with regard to affordance preservation. The data demonstrate that there are some convergencies (more than one-third of affordances are preserved by both, humans and LLMs), but there are also divergencies (close to one-third of affordances are preserved by only one system).

Table 7: Differences in affordance preservation.

	Affordance preservation
Both Human & LLM	35% (26)
Human only	12% (9)
LLM only	19% (14)
Neither	34% (25)

To assess how strongly metaphor preservation is associated with inferential affordance preservation in each system, humans and LLM, a logistic regression was performed.

For both humans and the LLM, metaphor preservation is a statistically significant predictor of affordance preservation ($p < .001$ in both instances). For humans, when the metaphor is removed, it is unlikely that affordance is preserved (18.18%, intercept -1.50, $p < .001$). Retaining the metaphor increases the likelihood that affordance will be preserved by 40.5x (95% confidence interval [9.8, 165], $z = 5.12$, $p = < 0.001$). For LLMs, when the metaphor is removed, it is also unlikely that affordance is preserved, but more likely than in the human examples (32.56%, intercept -0.73, $p = 0.025$). Metaphor preservation is a weaker predictor of affordance preservation for the LLMs but still increases likelihood by 10.77x (95% CI [3.41, 34.02], $z = 4.05$, $p < .001$). Both approaches show strong fit, although the human fit is stronger than the LLM fit (McFadden pseudo-R^2 of 0.40 and 0.20 respectively).

The logistic regression indicates that metaphor preservation is more strongly associated with affordance preservation for humans than for the LLM.

Discussion and Conclusions

The results of the two studies provide a novel perspective on the current theoretical debate about the role of metaphors in language and cognition. These findings are best understood through the construct of inferential affordance. This contrast allows us to distinguish between linguistic form (metaphor) and the inferential meaning licensed by that form (inferences about causal relations and structures). Focusing on surface representation alone (linguistic metaphor) obscures how different representational choices support or constrain inferential meaning.

We report two main findings. First, metaphors are not necessary to convey inferential affordances: inferential affordances are sometimes preserved even in cases when metaphors are removed. This demonstrates that representational strategy (metaphor) can be dissociated from inferential content. Second, for humans the association between metaphor as a representational strategy and inference is nevertheless rather strong. Although retention of inferential affordance is possible without the metaphor, when metaphors are dropped, inferential affordances are more often lost as well. This suggests that while there are different representational choices, humans strongly prefer metaphors as means to express inferential meaning.

Our findings have specific implications for the theories of metaphors. Under a strong version of CMT, metaphors structure conceptual understanding. Therefore, the survival of inferential affordances in case of metaphor removal is inconsistent with the strong CMT. However, it leaves open the possibility that metaphors play a role in conceptual organization. The finding that metaphors are optional rather than required for conveying inferential meaning aligns with the views expressed by Pinker (2007) and pragmatic approaches. However, the fact that humans strongly prefer metaphor as a vehicle for inferential meaning is difficult to reconcile with the view that metaphors are primarily rhetorical devices. Instead, our results align most closely with the view on metaphors articulated by Holyoak (2019), according to which metaphors are optional but can be selectively preserved when they support understanding.

In conclusion, we argue that metaphors should be understood neither as foundational cognitive mechanisms nor as mere rhetorical tools, but as preferred representational choices that humans employ to convey inferential meaning.

References

Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning through case comparisons: A meta-analytic review. *Educational Psychologist*, *48*(2), 87–113.

Duit, R. (1991). On the role of analogies and metaphors in learning science. *Science Education*, *75*(6), 649–672. https://doi.org/10 .1002/sce.3730750606

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7*, 155–170. http://dx.doi.org/10.1207/s15 516709cog0702_3

Gentner, D., & Clement, C. (1988). Evidence for relational selectivity in the interpretation of analogy and metaphor. In G. H. Bower (Ed.), *Advances in the psychology of learning and motivation* (Vol. 22, pp. 307–358). New York, NY: Academic Press. http://dx.doi.org/10.1016/ S0079-7421(08)60044-4.

Gibbs Jr, R. W., & O'Brien, J. E. (1990). Idioms and mental imagery: The metaphorical motivation for idiomatic meaning. *Cognition*, *36*(1), 35–68.

Gibbs, R. W., Jr., Bogdanovich, J. M., Sykes, J. R., & Barr, D. J. (1997). Metaphor in idiom comprehension. *Journal of Memory and Language*, *37*, 141–154. http://dx.doi.org/10.1006/jmla.1996.2506

Gibson, J. J. (1979). *The ecological approach to visual perception*. Houghton Mifflin.

Glucksberg, S., Brown, M., & McGlone, M. S. (1993). Conceptual metaphors are not automatically accessed during idiom comprehension. *Memory & Cognition*, *21*, 711–719. http://dx.doi.org/10.3758/BF031 97201

Gómez, J. C. (2022). Comparative psychology. In J. Vonk & T.K. Shackelford (Eds.), *Encyclopedia of animal cognition and behavior* (pp. 1569–1583). Springer International Publishing.

Gray, M. E., & Holyoak, K. J. (2021). Teaching by analogy: From theory to practice. *Mind, Brain, and Education*, *15*(3), 250–263.

Harris, R. J., Tebbe, M. R., Leka, G. E., Garcia, R. C., & Erramouspe, R. (1999). Monolingual and bilingual memory for English and Spanish metaphors and similes. *Metaphor and Symbol*, 14(1), 1–16.

Hicke, R. M. M., & Kristensen-McLachlan, R. D. (2024). Science is exploration: Computational frontiers for conceptual metaphor theory. *Proceedings of the Computational Humanities Research Conference*, pp 1–12.

Holyoak, K. J. (2019). *The spider's thread: Metaphor in mind, brain and poetry*. Cambridge, MA: MIT Press.

Ichien, N., Stamenković, D., & Holyoak, K. (2024). Interpretation of novel literary metaphors by humans and GPT-4. *Proceedings of the Annual Meeting of the Cognitive Science Society*.

Jiang, C., Maddela, M., Lan, W., Zhong, Y., & Xu, W. (2020). Neural CRF model for sentence alignment in text simplification. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 7943–7960.

Jorba, M. (2019). Husserlian horizons, cognitive affordances and motivating reasons for action. *Phenomenology and the Cognitive Sciences*, 5(5), 1–22.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago, IL: University of Chicago Press.

Low, G. (2008). Metaphor and education. In R. W. Gibbs, Jr., (Ed.), Cambridge handbook of metaphor and thought (pp. 212–231). Cambridge, UK: Cambridge University Press. http://dx.doi.org/10.1017/CB O9780511816802.014

McClelland, T. (2020). The mental affordance hypothesis. *Mind*, 129(514), 401–427.

McCoy, R. T., Yao, S., Friedman, D., Hardy, M. D., & Griffiths, T. L. (2024). Embers of autoregression show how large language models are shaped by the problem they are trained to solve. *Proceedings of the National Academy of Sciences*, *121*(41), e2322420121.

Murphy, G. L. (1996). On metaphoric representation. *Cognition*, *60*, 173– 204. http://dx.doi.org/10.1016/0010-0277(96)00711-1

Nayak, N. P., & Gibbs, R. W., Jr. (1990). Conceptual knowledge in the interpretation of idioms. *Journal of Experimental Psychology: General*, *119*, 315–330. http://dx.doi.org/10.1037/0096-3445.119.3.315

Pinker, S. (2007). *The stuff of thought*. New York: Viking.

Siddharthan, A. (2014). A survey of research on text simplification. *ITL-International Journal of Applied Linguistics*, *165*(2), 259–298.

Smirnova, A., Chun, K. B., Rothman, W. L., & Sarma, S. (2025a). Text simplification for children: Evaluating LLMs vis-à-vis human experts. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp. 1–10.

Smirnova, A., Lee, E. S., & Li, S. (2025b). Numeric information in elementary school texts generated by LLMs vs human experts. *Proceedings of the Artificial Intelligence in Measurement and Education Conference (AIME-Con),* pp. 183–191.

Sperber, D., & Wilson, D. (2008). A deflationary account of metaphors. In R. W. Gibbs (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 84–105). Cambridge: Cambridge University Press.

Steen, G. L. A. (2010). Method for linguistic metaphor identification: from MIP to MIPVU. John Benjamins Publishing.

Thorndike, E. L. (1911). *Animal intelligence*. New York: Macmillan.

Wolff, P., & Gentner, D. (2011). Structure-mapping in metaphor comprehension. *Cognitive Science*, *35*, 1456 – 1488.http://dx.doi.org/10.1111/j.1551–6709.2011.01194.x

Wolska, M., & Clausen, Y. (2017). Simplifying metaphorical language for young readers: A corpus study on news text. *Proceedings of the 12th Workshop on Innovative Use of NLP for Building Educational Applications* (pp. 313–318).

Xu, W., Callison-Burch, C., & Napoles, C. (2015). Problems in current text simplification research: New data can help. *Transactions of the Association for Computational Linguistics*, *3*, 283–297.