Looking at image schemas: Combinations and modifications

Michael J. Spivey and Teenie Matlock
Department of Cognitive and Information Sciences
University of California, Merced
5200 Lake Rd., Merced, CA 95343 USA

Abstract

The goal of this study was to revisit the nature and role of image schemata in language use. Image schemas have often been thought to be low-level quasi-primitives that structure much of language and thought (Johnson, 1987; Oakley 2010). Over the years, they have been studied by cognitive scientists who are interested in the semantics of language, both literal and non-literal. In this study, we queried naive participants' intuitions about various verbs in everyday English sentences. The events they denoted have various visuospatial orientations: horizontal, vertical, or some combination or modification of those orientations. The results of our survey, which follow on earlier work (Richardson, Spivey, Edelman & Naples, 2001), show robust consistency among people's intuitions and provide further insights into how image schemas work, in particular, how they are dynamic, flexible, and combine to create meaning.

Keywords: Image schemas, cognitive linguistics, embodied cognition, event representation, semantics.

Introduction

Everyday communication is motivated by our most basic bodily experiences with physical space, including how we stand, sit, see, breathe, push and pull objects, eat, and move from one point to another. Think about the sense of containment. It underlies phrases like "in this room", "in the box", "deep in my heart", and "full of anger". Or consider motion. Motion is inferred in statements like "sprinted across the finish line", "drove home", "headed for trouble", "going downhill", and "jump to the next topic".

For the past few decades, research in the broad area of embodied cognition suggests that image schemas, such as CONTAINMENT and PATH shape our everyday literal and non-literal thought and language. They are developed over time as we interact with and learn about the physical world (see Gibbs, 1996; Grady, 2005, Hampe, 2005; Johnson, 1987; Lakoff, 1990; Mandler, 1992). These implicit embodied elements are apparent in literal and non-literal language across cultures and situations, and as such, they help us understand what is new, complex, or abstract (Lakoff & Johnson, 1980). For instance, the CONTAINMENT image schema is prevalent in discourse about ingroups and outgroups and social exclusion in discussions of immigration (Charteris-Black, 2006). And they are part of grammatical structure (Fillmore & Kay, 1995; Goldberg, 1995; Langacker, 1987). Simply stated, using a transitive verb signals that some entity creates a state change in another entity (or itself), as in "They painted the wall", "She ripped the towel", and "Rob cut himself."

As events often take place along a path in both space and time, it should not be surprising that temporal duration is often described with metaphors about spatial extent (Matlock, Ramscar, & Boroditsky, 2005). Take, for example, when a person gestures toward the right to refer to the future portion of a timeline, or when they point behind themselves to reference the past. Torralbo, Santiago, and Lupiáñez (2006) showed evidence that speakers can switch between the lateral axis (left-right) or the sagittal axis (backward-forward) when speaking with gestural metaphors about time. Following up on that work, Walker and Cooperrider (2016) found evidence suggesting that speakers will sometimes combine the two spatial axes to produce a gesture that lies along the *diagonal* between them (see also, Marghetis, McComsey, & Cooperrider, 2020). These results suggest a continuous co-activation of multiple metaphors at once, rather than the selection of only one at a time. In the present work, we extend this metaphor-combination perspective to explore participants' intuitions about the horizontal or vertical spatial axes associated with a variety of common verbs. Like Walker and Cooperrider's study, we, too, find that people can naturally combine these two axes to entertain a diagonally-oriented image schema.

Previous work has shown that intuitions are consistent across people with regard to imagistic properties of metaphors (Gibbs & O'Brien, 1990), proverbs (Gibbs, Ström, & Spivey-Knowlton, 1997), and image schemas (Richardson, Spivey, Edelman & Naples, 2001). For example, Scheerer and Lyons (1957) asked participants to match the labels "gold", "silver", and "iron" to a sinewave, a sawtooth wave, and a square wave. Far from a random or idiosyncratic set of responses, most participants chose the same mapping. The smooth curves of the sinewave were seen as corresponding well to the gentle sheen of gold. The sharp sawtooth wave was seen as matching the harsher glare of silver. The square wave was seen as befitting the sturdy construction uses for iron.

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society.*

In a similar manner, Richardson, Spivey, Edelman and Naples (2001) gave participants four image schema response options to choose among (two horizontally oriented and two vertically oriented) to correspond to rebus sentences like "○ pushed □ " and "○ lifted □ ". For the *pushed* event, 88% of naive participants selected the horizontal image schema with a rightward-pointing arrow (see Figure 1, top left). For the *lifted* event, 87% of naive participants selected the vertical image schema with an upward-pointing arrow (see Figure 1, top right). On average, such high-concreteness verbs elicited 75% agreement on the specific image schema chosen, and 90% agreement on the axis chosen (horizontal or vertical). Although the verbs with low concreteness, such as "○ warned □ " and "○ respected □ ", elicited agreement well above chance performance, there was only 55% agreement on the specific image schema chosen, and 67% agreement on the axis chosen.

Richardson, Spivey, Barsalou, and McRae (2003) followed up that survey with reaction-time experiments that presented spoken sentences using those verbs, and found that concurrent visuospatial tasks were influenced by the orientation of the verb's image schema. The results were clearly more robust for the concrete verbs than for the abstract verbs. This observation was corroborated when Bergen, Lindsay, Matlock and Narayanan (2007) replicated the results only for the concrete verbs. Further work in this embodied cognition tradition has shown that concrete action verbs (but not nouns) influence the acceleration profiles of concurrent natural reaching movements (Nazir et al., 2008).

In our present experiment, we explored the possibility of these two-dimensional visuospatial image schemas being combined or modified in naive participants' intuitions about the shape and orientation of a *conjoined* event or a *grammatically reflexive* event. Take, for example, the following rebus sentence: "○ pushed and lifted □ ". If people can form a linear combination of their horizontal image schema for *pushed* and their vertical image schema for *lifted*, then perhaps they can entertain a diagonal image schema for "pushed and lifted" (see Figure 1, bottom). Similarly, if more abstract verbs are capable of this linear combination, then "○ warned and improved □ ", might allow people to compose a diagonal image schema from their horizontal image schema for *warned* and their vertical image schema for *improved* (see Figure 2).

Further modifications of image schemas might be possible in people's intuitions when considering one of these verbs in a grammatically reflexive rebus sentence, as in "○ and □ lifted each other" and perhaps also "○ and □ improved each other." Figures 3 and 4 depict this transformation of converting a vertical image schema into a horizontal image schema as a result of the grammatical reflexivity in the sentence.

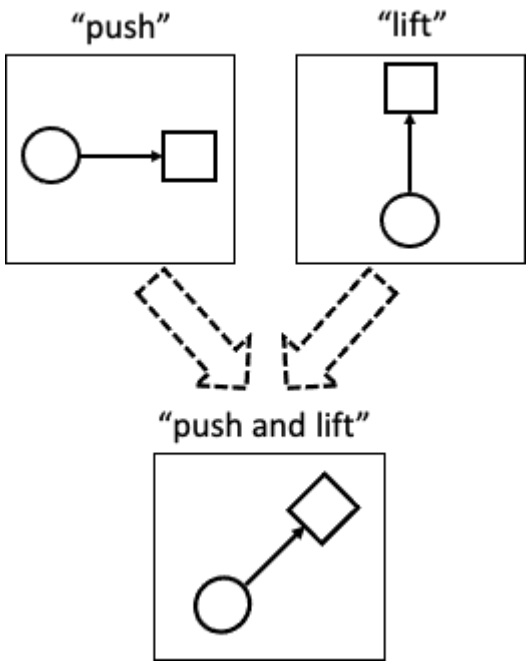


Figure 1: Conjunctions of spatially oriented concrete verbs may result in novel combinations of image schemas.

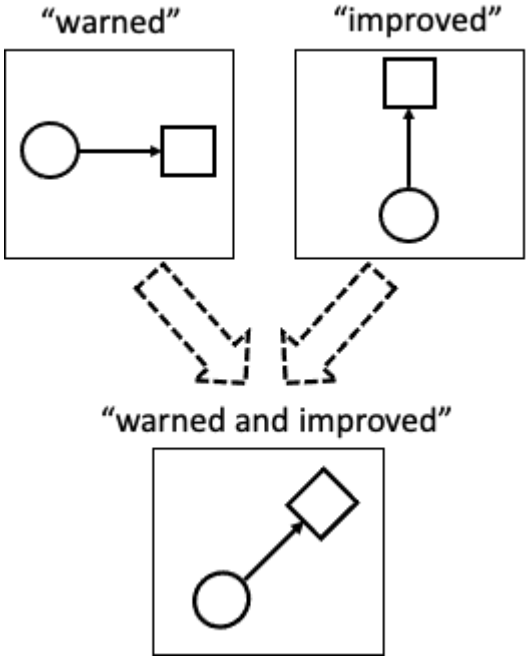


Figure 2: Conjunctions of spatially oriented abstract verbs may also result in novel combinations of image schemas.

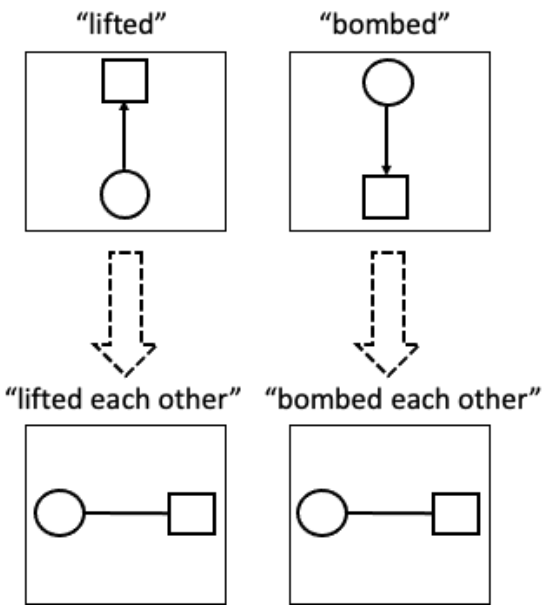


Figure 3: Reflexive constructions with concrete vertical verbs may result in horizontal non-directional image schemas because neither Subject is dominant when the grammar is reflexive.

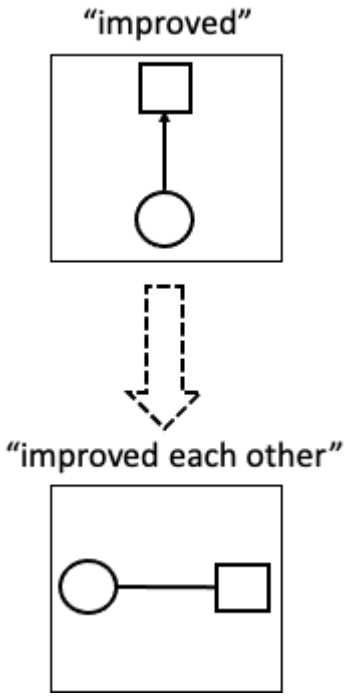


Figure 4: Reflexive constructions with abstract vertical verbs may also result in horizontal non-directional image schemas because neither Subject is dominant when the grammar is reflexive.

Methods

Participants

Fifty University of California, Merced students (ages 18-21; 32 female) in an Introduction to Cognitive Science course completed the image schema survey for extra credit. None had instruction in image schema theory.

Stimuli and Procedure

Fifteen rebus sentences were constructed involving a circle (as Subject), a simple past tense verb, and a square (as Object), drawing significantly from the stimuli of Richardson et al. (2001). An additional 5 rebus sentences were included that involved a circle and a square as the conjoined Subjects, a simple past tense verb and the phrase "each other" to make the constructions grammatically reflexive. See the leftmost portion of Figure 5. In contrast to the Richardson et al. (2001) survey, which included 4 image schema response options, the present survey provided 8 image schema response options to choose from (topmost row of Figure 5). In addition to Richardson et al.'s original four image schema options (C, E, F, and H in Figure 5), we added four new image schema options (A, B, D, and G in Figure 5). Image schema F naturally depicts the common rightward directionality of most action verbs in English and other languages that are read left-to-right (Chatterjee, 2001). Image schemas C, E, and H depict leftward, downward, and upward directionalities of events, respectively. The new image schema options that we added provided participants the opportunity to choose images that depict a *diagonal* down-and-rightward directionality of movement (A); a downward movement of the square toward the circle (B); a *diagonal* up-and-rightward directionality of movement (D); and a *non-directional* horizontal option (G). The new *diagonal* response options (A and D) were intended to accommodate any participants who might potentially combine rightward and vertical image schemas in their impressions of the event. The new non-directional response option (G) was intended to accommodate participants who would potentially convert what might be a vertical image schema into a horizontal one when the sentence involved a reflexive grammatical construction, such as "○ and □ lifted each other." (Note that image schema B was little more than a foil, and it was very rarely chosen.)

In one survey, the rebus sentences were listed on the right half of the page, while the 8 image schema response options were labeled A.-H. and arrayed in a 2X4 grid on the left half of the page. Participants were instructed to select one of the 8 diagrams that best depicts the rebus sentence, and were shown the example rebus sentence, "○ chased □ ", which was described as matching the rightward image schema (F). Next to each rebus sentence, participants wrote in the letter A to H to indicate which image schema they thought matched the event. In an

attempt to prevent participants from potentially perceiving the grid of image schemas as though from an aerial perspective (thus losing upward and downward connotations), some depth-perspective lines were added inside each box containing an image schema (see top row of Figure 5). Participants completed the survey in approximately 15 minutes.

Results

Far from random or idiosyncratic in their responses, participants showed substantial agreement in their choices of image schema for most of the rebus sentences. Each rebus sentence elicited a response distribution that significantly deviated from a flat distribution as measured by a Chi-square goodness-of-fit test (all *p*s< .05).

Several rebus sentences replicated the results from Richardson et al. (2001). For instance, the verb *pushed* elicited 82% agreement on the horizontal rightward image scheme (F, in Figure 5) and the verb *lifted* elicited 82% agreement on the vertical upward image schema (H, in Figure 5). Similar to the reduced agreement with abstract verbs found in Richardson et al., we also found that the verb *warned* elicited 42% agreement on the horizontal rightward image scheme (F, in Figure 5) and the verb *improved* elicited 42% agreement on the vertical upward image schema (H, in Figure 5). Interestingly, the abstract verb *respected* elicited a pattern of responses that was noticeably different from what it elicited in the Richardson et al. study. In the following sections, we examine: a) how conjoined verb phrases elicited selection of diagonal image schemas, as a composition of the horizontal and vertical images schemas associated with each individual verb, b) how reflexive syntax (where the circle and square are both doing the action to each other) converted vertically-oriented image schemas into horizontal ones, and c) the curious case of the verb *respected*.

Combining Image Schemas

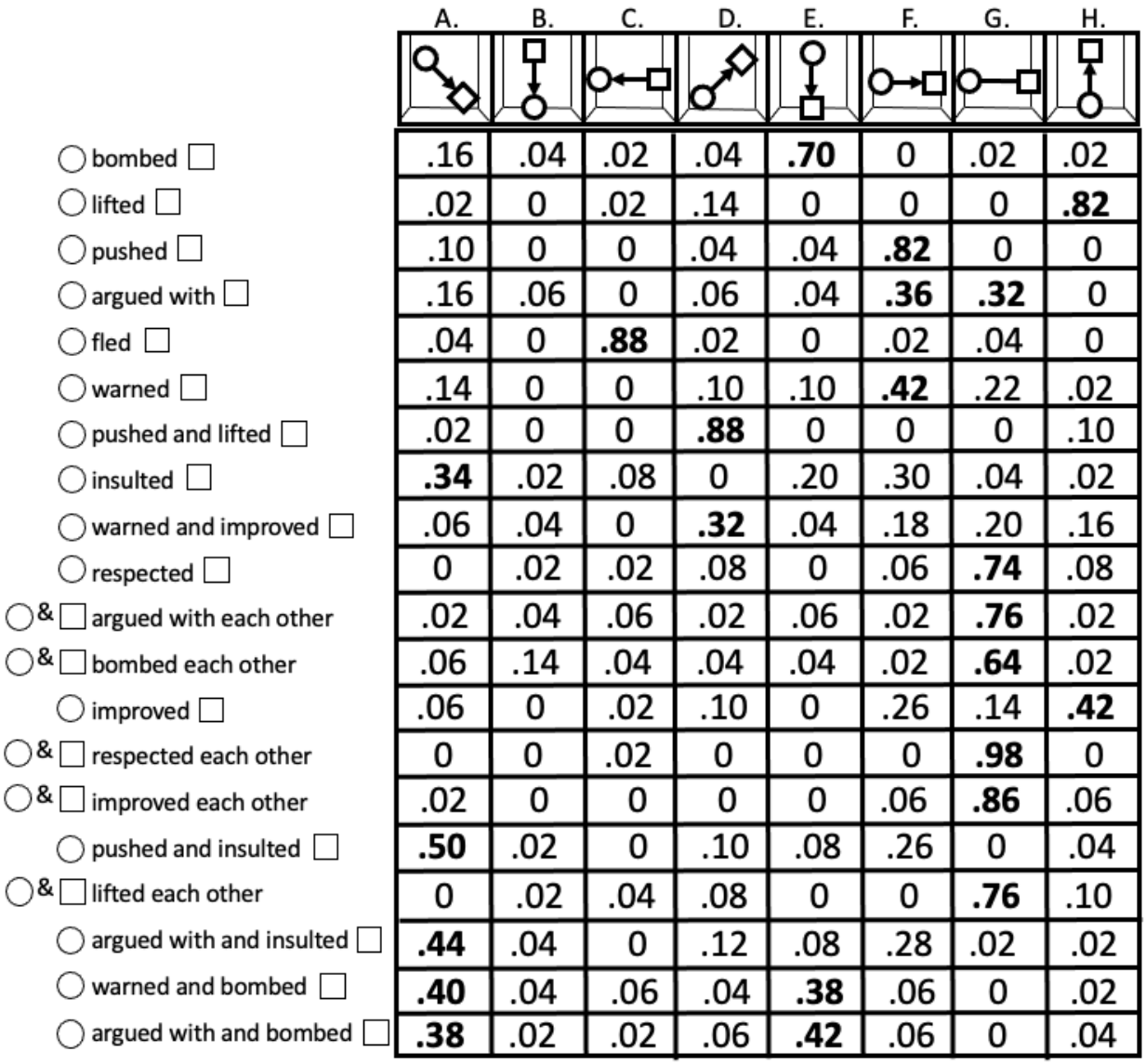
Several of the stimulus items in the survey tested whether naive participants would have intuitions about conjoined verb phrases that combined the two image schemas associated with those two verbs. For example, when the two concrete action verbs *pushed* and *lifted* were used in the rebus sentence "◯ pushed and lifted □ ", participants very consistently chose the diagonal upward-and-rightward image schema D 88% of the time (see Figure 5). When a concrete action verb was conjoined with a more abstract verb, "◯ pushed and insulted □ ", participants somewhat less-consistently chose the diagonal downward-and-rightward image schema A 50% of the time. When two abstract verbs were conjoined, as in the rebus sentence "◯ argued with and insulted □ ", we observed 44% agreement on the diagonal image schema A. Similar results with "◯ warned and improved □ " showed 32% agreement on the diagonal image schema D. The conjoined verb phrases of "warned and bombed" and "argued with and bombed" were less consistent, eliciting some ambivalence between the diagonal image schema A and the vertical image schema E.

Reflexive Syntax

Several of the stimulus items in the survey were used to test whether people would have intuitions about grammatically reflexive constructions that entail the action being carried out *by* both Subjects and *on* both Subjects, such as "◯ and □ bombed each other." In the case of a verb whose image schema is typically vertical, such as *bomb*, the implication is typically that one of the participants in the event is above the other. This is what makes the event take on a vertical orientation in its two-dimensional image schema layout. However, in the reflexive grammatical construction, each event participant is serving as both Agent and Recipient of the event. Therefore, a vertical arrangement of the two participants is less felicitous. In the case of the concrete action verb *bombed*, the transitive sentence, "◯ bombed □ ", consistently elicited the vertical downward image schema E 70% of the time (see Figure 5). However, when that same verb was used in a reflexive grammatical construction, "◯ and □ bombed each other", participants consistently chose the horizontal non-directional image schema G 64% of the time (Figure 5). Similarly, the verb *lifted* in its transitive form elicited 82% agreement on the vertical upward image schema H, but when placed in a reflexive grammatical construction, e.g., "◯ and □ lifted each other", suddenly the preferred image schema is the horizontal non-directional option G with 76% agreement (see Figure 5). This result was replicated with the more abstract verb *improve* as well. In its transitive form, *improve* elicited 42% agreement on the vertical upward image schema H, but when placed in a reflexive grammatical construction, "◯ and □ improved each other", the preferred image schema was the horizontal non-directional option G with 76% agreement (Figure 5).

The Curious Case of "Respect"

An interesting counterexample to this general pattern is the verb "respect." While "◯ and □ respected each other" elicited 98% agreement on the horizontal non-directional image schema (G), the transitive version of that sentence ("◯ respected □ ") was already eliciting 74% agreement on that same horizontal option anyway. In the original survey by Richardson et al. (2001), with the verb "respect," the vertical upward image schema (H, in Figure 5) garnered far more responses (54%) than the other three response options, C (14%), E (4%), and F (29%). Given that set of four options, it was the clear favorite. However, in the present study – with eight image schema options, including a non-directional one – the clear favorite was option G, the



	A.	B.	C.	D.	E.	F.	G.	H.
◯ bombed ☐	.16	.04	.02	.04	**.70**	0	.02	.02
◯ lifted ☐	.02	0	.02	.14	0	0	0	**.82**
◯ pushed ☐	.10	0	0	.04	.04	**.82**	0	0
◯ argued with ☐	.16	.06	0	.06	.04	**.36**	**.32**	0
◯ fled ☐	.04	0	**.88**	.02	0	.02	.04	0
◯ warned ☐	.14	0	0	.10	.10	**.42**	.22	.02
◯ pushed and lifted ☐	.02	0	0	**.88**	0	0	0	.10
◯ insulted ☐	**.34**	.02	.08	0	.20	.30	.04	.02
◯ warned and improved ☐	.06	.04	0	**.32**	.04	.18	.20	.16
◯ respected ☐	0	.02	.02	.08	0	.06	**.74**	.08
◯ & ☐ argued with each other	.02	.04	.06	.02	.06	.02	**.76**	.02
◯ & ☐ bombed each other	.06	.14	.04	.04	.04	.02	**.64**	.02
◯ improved ☐	.06	0	.02	.10	0	.26	.14	**.42**
◯ & ☐ respected each other	0	0	.02	0	0	0	**.98**	0
◯ & ☐ improved each other	.02	0	0	0	0	.06	**.86**	.06
◯ pushed and insulted ☐	**.50**	.02	0	.10	.08	.26	0	.04
◯ & ☐ lifted each other	0	.02	.04	.08	0	0	**.76**	.10
◯ argued with and insulted ☐	**.44**	.04	0	.12	.08	.28	.02	.02
◯ warned and bombed ☐	**.40**	.04	.06	.04	**.38**	.06	0	.02
◯ argued with and bombed ☐	**.38**	.02	.02	.06	**.42**	.06	0	.04

Figure 5: Proportions of selected image schemas for each rebus sentence, in the order presented in the survey. (Proportions >0.3 are in bold)

horizontal non-directional image schema, even when the verb was used transitively. As this contrasts with the vertically-oriented interpretation of "respect" observed by Richardson et al. (2001), we asked some of our participants to explain their choice of option G for the rebus sentence "◯ respected ☐." One participant admitted to waffling between choosing the vertical H option and the horizontal G option and eventually settled on G because "true respect usually has some mutual quality to it." Indeed, the fact that "respect" may not be as vertical-oriented as previously suspected may have contributed to the generally weaker effects of *abstract* verbs' image schemas influencing real-time processing (Bergen et al. 2007; Richardson et al., 2003).

Discussion

These results offer new insights into how people make sense of action descriptions, such as pushing, arguing, and fleeing. We focused on horizontal and vertical relationships in a two-dimensional image schematic framework. For instance, their thoughts about the action of fleeing involves a horizontal schema, one associated with horizontal movement away from another entity, such as a location. Similarly, their thoughts about bombing include a vertical schema, with downward movement. Richardson et al. (2001) found consistent judgments about these image schema orientations. Here, we expanded that work by looking at what would happen when participants

could opt for diagonal relationships. In doing so, we found that people appear to naturally combine image schematic intuitions, and this can lead to diagonal choices. The results were robust with high-concreteness verbs such as *push* and *lift*, but also present with low-concreteness verbs such as *argue* and *insult*. This ability to combine two image schemas to make a novel composite image schema suggests that the two-dimensional visuospatial layouts that people use for conceptualizing the space of events are flexible in ways that have not been documented before.

In addition to combining two image schemas, the results here suggest that people can also modify individual image schemas as a result of grammatical manipulations. When the verb and two entities are placed in a reflexive grammatical construction, e.g., "they [verb]ed each other", participants readily convert their vertical image schemas (for verbs like *bomb*, *lift*, and *improve*) into horizontal image schemas.

These results are promising and deserve further attention, for they show how the image schema formation process is flexible and innovative. Rather than being a fixed set of static immutable building blocks for cognition, image schemas are merged in real-time to make a new one and an image schema can also interface with grammatical constraints to make a new one. The new image schemas that are constructed on-the-fly for these conjoined-verb events and reflective-syntax events may or may not be of the same status as the original image schemas that were used to make these new ones. However, as a child acquires perceptual evidence corresponding to the most basic image schemas (Mandler, 1992; Rohrer, 2005), one can imagine combinations of those bases supporting the formation of new composite image schemas in slightly more complex events. For example, Mandler's proposed image-schematic concepts of ANIMACY and of CONTAINMENT (acquired by infants via perceptual experience) could perhaps be combined under certain experiential situations to form the slightly more complex new image-schematic concept of CONFINEMENT. With frequent enough exposure to circumstances that warrant the new image schemas, the newly formed ones could become stable internal representations of their own. The present work brings attention to this image schemata formation process and showcases its flexibility and innovation.

Future Work

We plan to conduct follow up experiments to strengthen and expand these initial findings, such as including multiple versions of the survey with the order of the sentences randomized. One task ahead of us will be to ensure that no participant would perceive the grid of image schemas as if they were taking an aerial perspective (thus, losing upward and downward connotations). Thus, in addition to depth-perspective lines inside the boxes that contain the image schemas, the example rebus sentence in the instruction portion could include a stickman labeled "You" standing up and facing the image schema and its box. This modification may increase the degree of agreement among participants for some of these verbs. Another extension of this work could be to test past progressive tense verbs, such as "◯ was improving □ ", especially given experimental findings that suggest that verbal aspect can affect event construal. For instance, in English the past progressive (which conveys imperfective aspect) can draw more attention to the sustained, ongoing nature of events than does the simple past, which conveys perfective aspect, placing emphasis on the completeness of the event (Anderson, Matlock, & Spivey, 2013; Matlock, 2011; see also Bergen & Wheeler, 2010).

References

Anderson, S. E., Matlock, T., & Spivey, M. (2013). Grammatical aspect and temporal distance in motion descriptions. *Frontiers in Psychology*, *4*, 337.

Bergen, B. K., Lindsay, S., Matlock, T., & Narayanan, S. (2007). Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cognitive Science*, *31*(5), 733-764.

Bergen, B., & Wheeler, K. (2010). Grammatical aspect and mental simulation. *Brain and Language, 112*(3), 150-158.

Charteris-Black, J. (2006). Britain as a container: Immigration metaphors in the 2005 election campaign. *Discourse & Society, 17*(5), 563-581.

Chatterjee, A. (2001). Language and space: Some interactions. *Trends in Cognitive Sciences, 5*(2), 55-61.

Fillmore, C. J., & Kay, P. (1995). Construction grammar. *Language*, *64*(501-538), 30.

Gibbs Jr, R. W. (1996). Why many concepts are metaphorical. *Cognition* , *61*(3), 309-319.

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 Jr, R. W. , Ström, L.K., & Spivey-Knowlton, M.J. (1997). Conceptual metaphors in mental imagery for proverbs. *Journal of Mental Imagery, 21*, 83-110.

Goldberg, A.E. (1995) *Constructions: A construction grammar approach to argument structure*. Chicago University Press.

Grady, J. E. (2005). Image schemas and perception: Refining a definition. In B. Hampe Beate & J. Grady (Eds.), *From perception to meaning: Image schemas in cognitive linguistics* (pp. 35-56). Berlin: Mouton de Gruyter.

Hampe, B. (2005). Image schemas in cognitive linguistics: Introduction. In B. Hampe Beate & J. Grady (Eds.), *From perception to meaning: Image schemas in cognitive linguistics* (pp. 1-12). Berlin: Mouton de Gruyter.

Johnson, M. (1987). *The body in the mind: The bodily basis of meaning, imagination, and reason*. Chicago: University of Chicago Press.
Lakoff, G. (1990). The invariance hypothesis: Is abstract reason based on image-schemas? *Cognitive Linguistics* 1(1), 39-74.
Lakoff, G., & Johnson, M. (1980.) *Metaphors we live by.* Chicago University Press.
Langacker, R.W. (1987/1991) Foundations of Cognitive Grammar (Vols I & II). Stanford University Press.
Mandler, J. M. (1992). How to build a baby: II. Conceptual primitives. *Psychological Review*, *99*(4), 587.
Marghetis, T., McComsey, M., & Cooperrider, K. (2020). Space in hand and mind: Gesture and spatial frames of reference in bilingual Mexico. *Cognitive Science*, *44*(12), e12920.
Matlock, T. (2011). The conceptual motivation of aspect. In K. Panther & G. Radden (Eds.), *Motivation in grammar and the lexicon* (pp. 133–147). Amsterdam: John Benjamins.
Matlock, T., Ramscar, M., & Boroditsky, L. (2005). On the experiential link between spatial and temporal language. *Cognitive Science*, *29*(4), 655-664.
Nazir, T. A., Boulenger, V., Roy, A., Silber, B., Jeannerod, M., & Paulignan, Y. (2008). Language-induced motor perturbations during the execution of a reaching movement. *Quarterly Journal of Experimental Psychology*, *61*(6), 933-943.
Oakley, T. (2007). Image schemas. In . Geeraerts & H. Vuyckens (Eds.), *The Oxford handbook of cognitive linguistics* (pp. 214-255). Oxford: Oxford University Press.
Richardson, D. C., Spivey, M. J., Barsalou, L. W., & McRae, K. (2003). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science*, *27*(5), 767-780.
Richardson, D. C., Spivey, M. J., Edelman, S., & Naples, A. J. (2001). Language is spatial: Experimental evidence for image schemas of concrete and abstract verbs. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (pp.845-850) Mahwah, NJ: Erlbaum.
Rohrer, T. (2005). Image schemata in the brain. In B. Hampe & J. Grady (Eds.), *From Perception to Meaning: Image Schemas in Cognitive Linguistics*. (pp. 165-195) Berlin: Mouton de Gruyter.
Scheerer, M., & Lyons, J. (1957). Line drawings and matching responses to words. *Journal of Personality, 25,* 251-273.
Torralbo, A., Santiago, J., & Lupiáñez, J. (2006). Flexible conceptual projection of time onto spatial frames of reference. *Cognitive Science*, *30*(4), 745-757.
Walker, E., & Cooperrider, K. (2016). The continuity of metaphor: Evidence from temporal gestures. *Cognitive Science*, *40*(2), 481-495.