Metaphors and Mind: An ERP Study of How the Brain Processes Metaphors

Crystal Poole, Tim Beyer, and David Andresen

University of Puget Sound
Abstract

Metaphors can differ in whether they are conventional (such as “love is war”) or novel (such as “love is a tidal wave”), and an unresolved question is if, and how, novel metaphors might become conventional through familiarity. To test this question, we asked participants to respond to experimenter-created literal phrases, conventional metaphors, novel metaphors, and participant-generated novel metaphors while measuring their brain activity. Previous studies have shown that a specific brain activity pattern, called the N400, reflects language processing such as whether a phrase is literal or metaphorical, and whether a metaphor is conventional or novel. We used the N400 to assess how different types of metaphors are processed, such as whether novel metaphors produced by the participants themselves are processed more like novel metaphors created by the researchers, or more like conventional metaphors. In addition, we correlated N400 patterns with individual characteristics such as creativity, crystallized intelligence (accumulated world knowledge), and fluid intelligence (problem-solving), to examine what cognitive abilities may be involved in processing different types of metaphors. Our study found that all metaphors, regardless of type, elicit similar N400 amplitudes and durations. Interestingly, after the metaphor creation task, conventional metaphors elicited a significantly greater N400 amplitude, more like novel metaphors. Participant-generated metaphors resembled conventional metaphors more than novel metaphors, indicating that conventionalization is related to the familiarity of a metaphor. Furthermore, few significant correlations were found between verbal cognitive abilities and metaphor processing, indicating that metaphor processing may require a unique verbal cognitive ability that was not tested.
METAPHORS AND MIND

Metaphors and Mind: An ERP Study of How the Brain Processes Metaphors

While literal language can be thought of as a sum of component meanings that are essentially “looked up” in memory and pieced together, non-literal language is much more difficult to understand for listeners and readers as, by nature, it involves language that does not make sense literally. A common form of non-literal language are metaphors. Metaphors equate two ideas to reveal a novel relationship between these ideas. For example, “he is a night owl” equates a person to an owl to reveal that like an owl, the person likes to be up at night. Metaphors demonstrate the complexity of language because they are by nature not true (he is not an owl), and yet we are able to interpret them with ease (like an owl, he is up at night). Despite being treated as specialty expressions only encountered in English classes, metaphors are extremely common in everyday language (Lakoff & Johnson, 1980). In The Metaphors We Live By, Lakoff and Johnson (1980) demonstrate that even idiomatic and other common language was based on metaphorical concepts, such as associating up with goodness in the phrase “on the up and up”. Therefore, studying how metaphors are understood in comparison to literal language both gives us insights into how more complex language is processed by the brain and how we conceptualize the world to construct meaning.

Metaphors

Metaphors differ in their familiarity. As such, metaphors are often categorized as conventional (common or familiar, such as “time is money”) and novel (uncommon or unfamiliar, such as “time is a whirlwind”). Bowdle and Gentner (2005) proposed the Career of Metaphor model, asserts that conventional and novel metaphors are processed using different cognitive processes. The Career of Metaphor model asserts that novel metaphors are processed through comparison, which is a process of comparing the semantic networks of the seemingly
unrelated ideas until some overlap is found to relate them. As an example, to process the novel metaphor “time is a whirlwind” the semantic network of time may conjure ideas like clocks, hours, minutes, passing, or flying, and the semantic network of a whirlwind might include air, spiraling, movement, and tornadoes among other things. By comparing the semantic networks of these two ideas, the concept of motion seems to be in both, therefore the metaphor would be taken to mean that time acts like a whirlwind in that it is constantly moving and is a force outside of a human control. Conventional metaphors, in contrast, are processed through categorization, which is the process of identifying a known semantic meaning and retrieving it from memory. For the metaphor “time is money”, instead of restarting the comparison process, which can be time consuming, the brain simply recognizes that it already has done the comparison many times and just retrieves the meaning from memory, understanding that time and money are both valuable resources.

Measuring Metaphor Processing in the Brain

There are many ways to study metaphor processing in the brain, but a variety of methodologies have indicated that there is at least some difference between the processing of literal phrases (e.g. “the time is 8 o’clock”) and metaphors (e.g. “time is a thief”), and many have also shown that novel and conventional metaphors are also processed differently (Bowdle & Gentner, 2005; Cardillo, et al., 2012; Goldstein, Arzouan, & Faust, 2012; Mashal & Faust, 2009). These studies are evidence for the Career of Metaphor model of metaphor processing proposed by Bowdle and Gentner (2005). However, the categories of novel and conventional are not as distinct as they seem because presumably, after repeated exposure or increased familiarity, novel metaphors must undergo some changes in processing to be processed like conventional metaphors, a process called conventionalization (Mashal & Faust, 2009). For example, Mashal
METAPHORS AND MIND

and Faust (2009) showed participants literal phrases, conventional metaphors, novel metaphors and unrelated word pairs while measuring the participants reaction time to whether the word pair they read formed a meaningful expression. They then re-exposed participants to the same set of stimuli and found that reaction times decreased across the two sessions. Furthermore, the reaction time was even shorter when novel metaphors were shown to the right hemisphere compared to the left in the first session, but there were not hemispheric differences in the second session. Mashal and Faust (2009) argue that the initially greater activity in the right hemisphere for novel metaphor processing, but lack of difference upon second exposure show evidence that the novel metaphors are processed differently during the second exposure. Therefore, the novel metaphors may begin conventionalization after only one exposure. Cardillo and colleagues (2012) showed participants three sets of novel metaphors at three levels of familiarity (new, 3 exposures, 5 exposures) while functional magnetic resonance images (fMRI) were taken. They found that the left inferior frontal gyrus (LIFG) and its right hemisphere homolog (RIFG) decreased in activity as familiarity increased. Both the LIFG and RIFG have both been found to increase activity when a participant must choose between competing semantic representations. Therefore, Cardio and colleagues (2012) argue that the decrease in activity in these areas as familiarity increases represents the shift from comparison to categorization that is proposed in the Bowdle and Gentner’s (2005) Career of Metaphor model. While these studies support that novel metaphors undergo some sort of conventionalization, both reaction times and fMRI data are not sensitive enough to the timing of brain activity to allow us to fully understand how metaphors are processed. However, event-related potentials (ERPs) can be used to record brain activity in real time.
METAPHORS AND MIND

Because ERPs, which detect brain waves at the scalp, measure brain activity in real time they are used to study how the brain constructs meaning, especially from language (Luck, 2000). ERPs are often described by components whose duration, amplitude, and timing are influenced by stimuli perceived by the brain. For example, the N400 is a component that is influenced by how easily we make sense of words we encounter (Kutas & Federmeier, 2000). Because the amplitude of ERP peaks can be thought of as representing the number of neurons acting in synchronization, a stimulus that creates a greater N400 peak amplitude is interpreted as requiring more processing resources (Luck, 2005). Greater amplitude and duration of the N400 component has been observed when words violate the expected meaning (e.g. “I like my coffee with cream and socks”), are not commonly used (e.g. ibex), or are more abstract (e.g. freedom) (Forgács, Bardolphp, Amsel, DeLong & Kutas, 2015; Kutas & Federmeier, 2000; Thibodeau & Durgin, 2011). How the N400 reflects the processing of metaphors is still debated; some researchers find a greater amplitude and duration (Kutas & Federmeier, 2000) while others point out the inconsistency of these findings (Brouwer & Crocker, 2017).

Because the N400 is sensitive to familiarity, an emerging pattern in research shows that literal phrases produce a low N400 amplitude, novel metaphors produce a large N400 amplitude, and conventional metaphors produce an amplitude in between the two (Goldstein et al., 2012). Goldstein et al. (2012) showed participants novel and conventional metaphors and asked participants to explain half of them. The participants were then shown the same novel and conventional metaphors again. They found that the N400 amplitude of explained novel metaphors was smaller than the amplitude of unexplained novel metaphors, while explained conventional metaphors had greater N400 amplitudes compared to unexplained conventional metaphors. These findings indicate that the brain’s strategy for interpreting novel and
METAPHORS AND MIND

conventional metaphors can change very quickly. Therefore, the brain must choose strategies for interpreting language stimuli based on recent exposure and familiarity.

Individual Differences in Metaphor Processing

Another approach to better understanding how conventional and novel metaphors are processed is by testing whether the same cognitive abilities are used during the processing of each metaphor type. Prior research has found that individual differences can impact metaphor production and comprehension (Beaty & Silvia, 2013; Gold, Faust, Ben-Artzi, 2012). In particular, verbal cognitive abilities, such as fluid intelligence (the ability to solve problems without prior information about the problem), crystallized intelligence (accumulated world knowledge), creativity, and retrieval ability (how quickly memories can be retrieved) have been linked to metaphor processing speed. Beaty and Silvia (2013) asked participants to produce novel and conventional metaphors and correlated verbal cognitive abilities with the quality of participant-generated metaphors. Results showed that the creation of novel metaphors is influenced by fluid intelligence and creativity, while the retrieval and use of conventional metaphors are influenced by crystallized intelligence and retrieval ability (Beaty & Silvia, 2013). Gold et al. (2012), found that people with higher verbal creativity scores had faster reaction times compared to those with lower scores when processing both novel and conventional metaphors. While reaction times may indicate some aspects of metaphor processing, metaphor processing is more complex than speed alone as the brain could be acting in a very different way, but still process the phrase in the same amount of time. Therefore, it is important to understand whether individual differences in verbal cognitive abilities correspond to brain activity during metaphor processing.

The Present Study
METAPHORS AND MIND

This study will compare brain activity (measured using ERP) related to the processing of conventional and novel metaphors to test how these metaphor types are processed differently. Moreover, we will investigate how processing of each of these metaphor types may change over time. Another purpose of this study is to understand whether familiarity mediates the difference in metaphor processing between metaphor types. To test this, we will ask participants to create novel metaphors and measure the N400 component in response to these metaphors at a later time. No studies have investigated how novel metaphors generated by a participant themself are processed at a later point in time by that same participant. Because the participants created the metaphors, they should be more familiar than other novel metaphors, but less familiar than conventional metaphors. This design will allow us to more definitely assess the process of conventionalization at an individual level. Finally, to our knowledge, no ERP study has investigated how individual differences in verbal cognitive abilities impact the N400 when encountering novel and conventional metaphors. Thus, we propose to link individual differences in verbal creativity, fluid intelligence, crystallized intelligence, and retrieval ability to properties of the N400 component while processing novel and conventional metaphors.

To investigate how conventional and novel metaphors are processed and how novel metaphors might be conventionalized, we conducted testing over three consecutive days. On Day 1, we measured participants’ brainwaves as they read and processed literal phrases, novel metaphors, and conventional metaphors (created by the experimenter). On Day 2, participants produced novel metaphors and took a series of offline measures of verbal cognitive abilities (retrieval ability, fluid intelligence, crystallized intelligence, and creativity; Burack, 1950; Kliegel, Martin, & Jäger, 2007; Mednick and Mednick, 1967; U.S. Army, 1942). On Day 3, ERP data was collected for all of the types of phrases from Day 1 in addition to the participant-
METAPHORS AND MIND

generated novel metaphors. We expect the Day 1 stimuli to assume the pattern previously observed by Goldstein et al. (2012), where the N400 amplitudes provoked by literal phrases are smaller than those of conventional metaphors, which are smaller than the N400 amplitudes related to novel metaphor processing. This pattern should be stable across time because participants are not shown the same metaphors twice. If familiarity is the primary factor that defines conventional metaphors, then the N400 from the participant created novel metaphors should pattern similarly to conventional metaphors. Following the results of Beaty and Silvia (2013), we expect fluid intelligence and creativity to be correlated with novel metaphor N400 amplitude, and crystallized intelligence to be correlated with conventional metaphor N400 amplitude.

Method

Participants

Twenty healthy participants from the University of Puget Sound volunteered for this study. Four participants were excluded from the study due to technological issues (n = 1) and excessive noise in the data (n = 3). The remaining sixteen participants (3 male, 11 female, 2 non-binary/non-conforming; mean age 21.00) all reported being fluent English speakers; fourteen of the participants reported exposure to at least one other language and eight participants identified themselves as bilingual. Participants were compensated $10 per session ($30 for all three sessions) and each session lasted no longer than an hour.

Materials

ERP data was collected using a tight-fitting electrode cap that detect electrical signals at the scalp; these electrical signals were converted into ERP components by averaging the electrical signals across many trials. A 32 channel Biosemi System referenced to the left and right mastoid electrodes and LabView were used to collect ERP data. MATLAB was used to
show the stimuli on a desktop screen in a lab setting. Participants could sit at a comfortable distance from the screen.

In this study, 80 conventional metaphors, 80 novel metaphors, and 80 literal phrases were used for the ERP stimuli (see Appendix for sample stimuli). The conventional metaphors came from the materials that had above average conventionality ratings from Roncero and de Almeida (2015), in addition, conventional metaphors were collected using the method described in Roncero and de Almeida (2015) using internet lists of common metaphor examples. The conventional metaphors provided the phrase stem and novel metaphors and literal phrases were created by attaching new words to the stem (e.g. *Time is money* $\rightarrow$ _______ *is money*). Words were added to the phrase stems to create the literal phrases (e.g. *A dollar is money*) and novel metaphors (e.g. *Education is money*). All of the phrases were also limited to be between 3 and 6 words long. Where possible, words were used from Cardillo, Watson, and Chatterjee (2017), that were normed for concreteness, but the list was too limited to use to create novel metaphors that made sense for all of the stimuli. The system of using stem phrases was adopted in order to create greater control for what words were in the phrases. To control for grammatical differences, only an *A-is-B* word structure was used (i.e. *A temper is a volcano*).

The offline measures were chosen to be language-specific and are based on the measures used by Beaty and Silvia (2013). The fluid intelligence, crystallized intelligence, and creativity measures were all shortened due to time restraints. Creativity was tested using the 15 odd-numbered problems from the Remote Associates Test (RAT) (Mednick & Mednick, 1967). Crystallized intelligence was tested using 25 questions from a multiple-choice vocabulary test selected using a random number generator (U.S. Army, 1942). Fluid intelligence was measured using the 10 odd-numbered problems from the letter sets task (Burack, 1950). Retrieval ability
METAPHORS AND MIND

was measured using the verbal fluency test from the Cognitive Telephone Screening Instrument (COGTEL) (Kliegel, Martin, & Jäger, 2007; see Appendix for samples of all measures). While some of these measurements were on the older side, they were chosen to correspond to the Beaty and Silvia (2013) study as well as target language-specific abilities.

Procedure

Participants were tested on three consecutive days and received $10 at the end of each testing session. On Day 1, after providing consent, participants’ brain activity was measured using ERP. While recording ERPs, participants were shown 120 stimuli (40 literal phrases, 40 conventional metaphors, 40 novel metaphors created by the experimenter) which were randomly sorted into 8 blocks per phrase type. After reading each stimulus participants were asked to press a key on a keyboard to indicate whether the phrase “makes sense” or not, or if they could imagine a context in which the phrase could “make sense”.

On Day 2, participants were asked to generate 40 novel metaphors by filling in the first word in a phrase (e.g. ________ is a journey) to create metaphors that are unfamiliar to them but make sense. The researcher first explained the task to participants and then gave the participants feedback on a practice stem and answered any of their questions about how to complete the task. Participants then completed all 40 metaphor stems while the researcher monitored their progress using a dual screen (which participants were informed of). The researcher took notes of any phrases produced by the participant that were either not a metaphor (too literal), not grammatically correct, or were on the list of experimenter-created metaphors. When the participant was finished creating metaphors they were asked to redo the metaphors that had been flagged by the researcher. Next, participants completed a demographics questionnaire
METAPHORS AND MIND

and several timed tasks to assess their intelligence (crystallized, fluid), retrieval ability, and creativity. All testing on Day 2 used Inquisit (Inquisit 5).

On Day 3, the same procedure as was used on Day 1 was employed, except participants responded to 160 phrases (40 literal phrases, 40 conventional metaphors, 40 novel metaphors created by the experimenter, 40 novel metaphors created by the participant on Day 2). The literal phrases, conventional metaphors, and novel metaphors created by the experimenter were counterbalanced between Day 1 and Day 3. Participants were then verbally debriefed.

Results

Raw ERP data were filtered in EMSE using a 12 db/octave infinite impulse response (IIR) bandpass filter with cutoffs at 1 and 20 Hz applied to all electrode channels (EMSE). The filtered data were then cleaned by selecting trials that did not have long periods of data (typically more than 100 ms, although slight adjustments were made depending on each participants’ data) that exceeded a 25 Hz deflection in either a positive or negative direction. The 300-500 ms time window was used to analyze the N400 peak, which was time-locked to the presentation of the final word in the phrase. Following conventions about where the N400 peak occurs, we restricted our analyses to the Cz electrode. The N400 peaks were considered to be the greatest peak during the 300-500 ms time window. The greatest amplitude during this time window was considered to be the N400 peak amplitude. The duration of the N400 peak was calculated by subtracting the two times when the peak that was considered the N400 peak amplitude crossed the x-axis. An alpha level of .05 was used for all analyses.

Processing Literal Phrases and Experimenter-Generated Metaphors Across Days

N400 peak amplitude. N400 peak amplitude has previously been shown to be different for literal phrases versus metaphors. The larger peak amplitudes that are produced by metaphors,
and particularly novel metaphors, represent greater amounts of neurons producing synchronized activity (Luck, 2005). In order to investigate whether peak amplitude differed by phrase type and day, a two-way repeated measures ANOVA that compared phrase type (literal phrase, conventional metaphor, novel metaphor) by day (Day 1, Day 3) was completed. There was a main effect of day on N400 peak amplitude, $F_{(1, 15)} = 6.02, p = .025$. There was no main effect of phrase type, $F_{(2, 30)} = 3.02, p = .064$. However, there was a significant interaction between phrase type and day, $F_{(2, 30)} = 3.64, p = .038$. To investigate the interaction effect between phrase type and day, paired sample $t$ tests were used, which indicated that conventional metaphor and novel metaphor peak amplitudes did not differ significantly ($t_{(15)} = -0.74, p = .473$). But, literal phrase peak amplitude was significantly less negative than conventional metaphor ($t_{(15)} = 3.09, p = .007$) and novel metaphor ($t_{(15)} = 3.76, p = .002$) peak amplitude (see Figure 1). Each phrase type was compared across days (e.g. conventional metaphors on Day 1 and conventional metaphors on Day 3) using paired sample $t$ tests, which showed that novel metaphors did not have significantly different peak amplitudes across days ($t_{(15)} = -1.16, p = .265$), literal phrases were not significantly different across days ($t_{(15)} = 0.31, p = .764$), but conventional metaphors had significantly greater N400 amplitudes on Day 3 compared to Day 1 ($t_{(15)} = 2.15, p = .048$; see Figure 2). Overall, the metaphor types produced similar N400 peak amplitudes, which were different from the literal phrases, but conventional metaphors produced at larger N400 amplitude on Day 3 than on Day 1.

**N400 peak duration.** Peak duration represents how long the brain spends on a particular stage of processing. Because the N400 has been shown to change depending on semantic content, the duration of an N400 peak can be thought of as representing the time it takes for the semantic content of a phrase to be processed. In the present study, to investigate whether peak
duration differed by phrase type and day, a two-way repeated measures ANOVA that compared phrase type by day was completed. There was a significant main effect of phrase type $F(2, 30) = 11.83, p = .025$. There was no main effect of day ($F(1,15) = 1.98, p = .180$) or an interaction ($F(2,30) = 0.90, p = .419$) for N400 duration. Paired sample $t$ tests showed that the two metaphor types, conventional metaphors and novel metaphors did not have significantly different durations ($t_{(15)} = 0.52, p = .608$). While literal phrases had a significantly shorter duration than both conventional metaphors ($t_{(15)} = 3.03, p = .008$) and novel metaphors ($t_{(15)} = 3.58, p = .003$). As is shown in Figure 3, the N400 peak duration for all metaphor types were similar and were greater than the peak duration of literal phrases, indicating that it takes longer to process the meaning of literal phrases compared to metaphors.

**Participant-Generated Metaphors Compared to Experimenter-Generated Phrase Types**

When comparing the participant-generated novel metaphors to other metaphor types, the peak amplitudes for literal phrases and novel metaphors were averaged across days and paired sample $t$ tests were used to analyze differences. As can be seen in Figure 4, participant-generated metaphor peak amplitude did not differ significantly from novel metaphors ($t_{(15)} = -0.13, p = .897$). However, the participant-generated metaphors were significantly more negative than the literal phrases ($t_{(15)} = -2.73, p = .015$).

Because peak amplitude significantly differed by day for conventional metaphors and therefore could not be averaged across days, paired sample $t$ tests were used to analyze the differences between participant-generate and conventional metaphor peak amplitude for each day that data was collected. The participant-generated metaphors were not significantly different from the conventional metaphors on Day 1 ($t_{(15)} = 0.28, p = .782$), but were significantly less negative than the conventional metaphors on Day 3 ($t_{(15)} = -2.80, p = .014$). So, for N400
amplitude, participant-generated metaphors resembled other metaphor types in general but were more similar to the conventional metaphors on Day 1 than on Day 3. As shown in Figure 4, the duration of participant-generated metaphors also did not differ significantly from conventional metaphors ($t_{(15)} = 0.90, p = .382$) or novel metaphors ($t_{(15)} = 0.08, p = .935$), but, the duration of the N400 for literal phrases was significantly shorter than the duration of participant-generated metaphors ($t_{(15)} = 3.88, p = .001$). The duration of participant-generated metaphors closely resembles the duration of other metaphor types, but not literal phrases.

**Individual Differences While Processing Metaphors**

We investigated whether participants’ scores on the verbal cognitive abilities measures were related to the N400 peaks that corresponded to their processing of metaphors and literal phrases. We correlated verbal cognitive abilities (retrieval ability, crystallized intelligence, fluid intelligence, creativity) with both N400 peak amplitude and duration. Surprisingly, there were no significant correlations between verbal cognitive abilities and phrase processing, except that crystallized intelligence had a significant negative correlation with conventional metaphor N400 duration ($r_{(16)} = -0.569, p = .021$). This indicates that as crystallized intelligence increased, less time was devoted to processing conventional metaphors. All correlation coefficients are shown in Tables 1 and 2. Surprisingly, we did not find evidence of individual variables correlating with verbal cognitive abilities despite the participants’ scores in each of the verbal cognitive abilities representing a wide range of scores.

**Discussion**

This study investigated how different types of metaphors (conventional, novel, and participant-generated) are processed in comparison to each other and literal phrases. Specifically, this study investigated whether the differences between conventional and novel metaphor processing are due to familiarity, and whether those processing differences would map on to
cognitive abilities used for language processing. This question was investigated by having participants generate novel metaphors, which were more familiar to them than the experimenter-created novel metaphors, but not like conventional metaphors, which are often familiar through years of exposure. Although we hypothesized that novel metaphors would have a significantly greater amplitude and duration, overall, our results show that metaphors, regardless of type, are similar in peak amplitude and duration for the N400. All metaphor types produce greater N400 amplitudes and durations as compared to literal phrases. So even though metaphors may vary in familiarity, with conventional metaphors being more familiar than novel metaphors, our findings strongly suggest that they still require similar language processing resources. However, while the metaphor types tested here appear to be processed similarly, they do not show identical processing patterns.

Unlike Goldstein et al. (2012), we did not find significant differences between conventional and novel metaphor processing. However, visual inspection of our data suggests that conventional metaphors have smaller, non-significant, peak amplitude than novel metaphors. This suggests that while very similar, conventional and novel metaphor processing may be slightly different, which would support the existing literature that assumes the two are processed differently.

**Conventional Metaphors Are Processed Differently Based on Context**

Although we hypothesized that amplitude and duration would remain consistent across days, conventional metaphor peak amplitude was significantly greater on Day 3 compared to Day 1. It appears unlikely that exposure to metaphors on Day 1 and Day 2 could have impacted how conventional metaphors are processed on Day 3. For example, Blasko and Connine (1993) used words from metaphorical phrases to try to prime their participants to process metaphors
faster. They found no difference in reaction time when participants were exposed to either a part of the metaphor or were shown words that were unrelated to the metaphor, indicating that exposure to words from a metaphor does not change metaphor processing. Therefore, even if the metaphor creation task on Day 2 had somehow primed participants by exposing them to words that were in metaphors they would later be exposed to, they should not have been primed to process the metaphors differently. Furthermore, the pattern observed in the present study is opposite from the pattern observed by previous research such that the peak amplitude increased across sessions, indicating that the change is not due to priming.

Although the increase in conventional metaphor amplitude cannot be explained by priming from the metaphor creation task, the metaphor creation task is still likely to have influenced the processing difference on Day 3. Because the stimuli were counterbalanced, this effect could not be due to the quality of the stimuli on each day since every participant saw a different combination of stimuli. Goldstein and colleagues (2012) found that when participants were asked to explain the meaning of conventional metaphors, the N400 amplitude increased. Therefore, the increase in conventional metaphor amplitude on Day 3 could indicate that participants were processing the conventional metaphors as though they were novel by focusing on re-explaining the meaning of a phrase they were already familiar with. This could be related to the phenomenon of hearing a very common saying and overthinking its meaning. This result indicates that though differences are not always found between novel and conventional metaphors they may just be too slight to detect with the current technology and may be heavily influenced by methodology. By showing that conventional metaphors can essentially be processed like novel metaphors if a participant is prompted to do so, this study illustrates that language processing is not fixed in the brain and while the brain may create heuristics to interpret
METAPHORS AND MIND

language (like conventional metaphors), once the heuristic is created the type of processing that is needed is decided situationally. Notably, the changes in language-processing mindset that likely occurred on Day 2 of the study lasted at least a full day, as they were not measured until Day 3. This indicates that the effect of changing the brain’s processing path are relatively long-lasting. Because this finding has never been observed before, further research needs to be conducted in order to understand why conventional metaphor processing might change over time.

Participant-Generated Novel Metaphors Are Conventionalized

Another piece of evidence that supports the idea that metaphor types might be processed differently is our finding that participant-generated novel metaphors have amplitudes that are more similar to conventional metaphors on Day 1 than on Day 3. We predicted that participant-generated metaphors would be processed more similarly to conventional metaphors than to novel metaphors because they would be more familiar than experimenter-created novel metaphors. Because conventional metaphors seemed to be processed more like novel metaphors on Day 3, the Day 1 conventional metaphor ERP patterns are likely a better representation of how conventional metaphors are typically processed. Therefore, because participant-generated metaphors produced ERP amplitudes that were not significantly different from the amplitudes produced by conventional metaphors on Day 1, participant-generated novel metaphors might be processed more like conventional metaphors. This supports the idea that novel metaphors are conventionalized through familiarity. Previous literature has not looked at how participant-generated metaphors are processed, but because they seem to behave more similarly to conventional metaphors than novel metaphors, they are likely processed similarly like conventional metaphors (through categorization). Furthermore, the brain must at some point
METAPHORS AND MIND

recognize that a metaphor is familiar and switch processing strategies. Because the duration of processing does not differ between metaphor types (conventional, novel, and participant-generated) conventional metaphor processing may not actually be faster, but might require fewer neuronal resources, hence the smaller N400 peak amplitude. Therefore, the brain must detect that a stimulus is familiar and change the processing strategy accordingly, which is why after just one exposure, participant-generated metaphors were processed like conventional metaphors.

Rethinking the Career of Metaphor Model

Our results provide mixed evidence for the Career of Metaphor model, which asserts that conventional and novel metaphors are processed differently, through the processes of categorization and comparison, respectively (Bowdle & Gentner, 2005). Our findings do indicate that there are two different types of metaphor processing, one that is typical of conventional metaphors, and one that is typical of novel metaphors, which supports the theory. Furthermore, the Career of Metaphor model accounts for increased exposure to a once novel metaphor (e.g. participant-generated novel metaphors) leading to conventionalization through categorization and the activation of memory. However, the Career of Metaphor model does not suggest that conventional metaphors could ever be processed like novel metaphors, relying on comparison during processing. Thus, the finding that conventional metaphors can be processed differently over time such that their N400 amplitudes increase problematizes the Career of Metaphor model for metaphor processing (Bowdle & Gentner, 2005). Given that this is not the first study to observe changes in conventional metaphor processing it is possible that conventional metaphors may change their processing path more readily than previously thought. However, this phenomenon should be investigated further given that this study is the first to test how different metaphors are processed over time. This study provides evidence that while conventional
METAPHORS AND MIND

metaphors typically follow a different processing trajectory than other metaphors, that processing pathway is not the only possibility.

Adding to the *Career of Metaphor* model, we propose that metaphors are processed through either categorization or comparison depending on the context of the situation and the familiarity of the metaphor. Processing through categorization is only possible when the metaphor is familiar, so the brain will check to see if a meaning can simply be pulled from memory. If the metaphor is not familiar, then the brain must begin the work of comparing the semantic networks of the two ideas in the metaphor until a meaning is found. The brain will tend to default to the strategy that it used most recently, in which case conventional metaphors may be processed through comparison when there have been recent high rates of exposure to unfamiliar metaphors. The brain will usually default to categorization, though, because that strategy would be more effective in understanding most of the language that we encounter as it is often highly familiar to us. This change to the *Career of Metaphor* model, might better account for the findings of this study by allowing for conventional metaphors to be processed either through categorization or comparison depending on the context. Further research needs to be done to test these changes to the *Career of Metaphor* model, or to reaffirm the existing model.

**Individual Differences on Metaphor Processing**

We predicted, based on previous research, that fluid intelligence and creativity would be related to novel metaphor processing, while crystallized intelligence would be related to conventional metaphor processing. However, individual differences in verbal cognitive abilities did not have much of an effect on the duration or peak amplitude of N400 components of metaphors. The only significant finding was that conventional metaphor duration was shorter when crystallized intelligence was greater. This is consistent with the findings of Beaty and
METAPHORS AND MIND

Siliva (2013) showed that the number of conventional metaphors participants could produce from memory based on a prompt was positively correlated with crystallized intelligence. This essentially indicates that the greater a person’s vocabulary and knowledge of metaphors, the faster they will process conventional metaphors. Generally, people who read more will have larger vocabularies and they may also be exposed to more conventional metaphors and get faster at recognizing and accessing the meaning of the metaphor from their memory (via categorization). This pattern is likely not found with novel metaphors because novel metaphors are ones that are unfamiliar and therefore will not be in memory, which is why they must be processed through comparison, as is stated by the Career of Metaphor model (Bowdle & Gentner, 2005). The absence of more significant correlations between the verbal cognitive abilities tested and metaphor processing may indicate that metaphor processing requires its own unique abilities, or the abilities that were tested are more generally important for language processing.

Conclusion

Future research must continue to parse out the differences between novel and conventional metaphors by using different tasks to elicit smaller and larger N400 amplitudes in conventional metaphors. While the metaphor creation task seems to have affected how conventional metaphors were processed, this assumption should be explored further. Further investigations of this phenomenon would allow researchers to better understand how language can be processed differently depending on the context in which it is presented. Future research may also compare responses to other participant-generated phrase types and phrases that participants are exposed to multiple times to determine whether participants respond differently.
METAPHORS AND MIND

to phrases they create compared to phrases that they have seen before. This would support that familiarity alone decreases N400 amplitudes.

A limitation of this study was that participant-generated metaphors could not be averaged from multiple testing sessions and compared across time. Further studies might show participant-generated metaphors multiple times, or after a longer time period to understand whether their conventionalization is maintained over time, becomes stronger over time, or only occurs within a certain time period. Another limitation of this study was that the metaphor creation task was not anticipated to affect metaphor processing, therefore it is unclear whether the N400 amplitude for conventional metaphors was greater on Day 3 because participants completed the metaphor creation task, or if conventional metaphors elicit larger amplitudes with increased exposure, or some other explanation. Further research should show participants unique literal phrases, conventional metaphors, and novel metaphors over an extended time-period (e.g. consecutive days), and with a greater break between testing sessions (e.g. a week between sessions) to better understand how the brain responds to repeated exposure to the same types of stimuli.

Overall, this study indicates that while metaphors are generally processed similarly to each other, there seem to be some distinct differences between how conventional and novel metaphors are processed that should be explored in future research. N400 amplitude seems to reflect the familiarity of a phrase, but duration of the N400 component does not differ between any metaphor types, indicating that changes in processing technique for conventional metaphors does not cause the metaphors to be processed faster, but that they require fewer cognitive resources to process them. Furthermore, novel metaphors can be processed more like conventional metaphors by increasing their familiarity, while conventional metaphors are processed like novel metaphors depending on context. This indicates that the brain is able to
change processing strategies, to view an already understood phrase in a new way. These properties indicate the flexibility of thought which allows ideas to be reimagined. This might explain why metaphors are so important to both daily life and poetry, as they allow new meanings to be created and old material to be reinvented.
References


<table>
<thead>
<tr>
<th>Cognitive ability</th>
<th>LP Amp</th>
<th>CM Amp</th>
<th>NM Amp</th>
<th>PG Amp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retrieval ability</td>
<td>0.434</td>
<td>-0.095</td>
<td>0.221</td>
<td>-0.044</td>
</tr>
<tr>
<td>Crystallized intelligence</td>
<td>-0.422</td>
<td>-0.276</td>
<td>-0.085</td>
<td>0.164</td>
</tr>
<tr>
<td>Fluid intelligence</td>
<td>-0.270</td>
<td>0.008</td>
<td>-0.063</td>
<td>-0.062</td>
</tr>
<tr>
<td>Creativity</td>
<td>-0.312</td>
<td>0.197</td>
<td>-0.185</td>
<td>0.003</td>
</tr>
</tbody>
</table>

*Note.* Amp = peak amplitude; Dur = duration; LP = literal phrases; CM = conventional metaphors; NM = novel metaphors; PG = participant generated

* *p < .05*
Table 2
Correlation matrix between verbal cognitive abilities and duration by phrase type

<table>
<thead>
<tr>
<th>Cognitive ability</th>
<th>LP Dur</th>
<th>CM Dur</th>
<th>NM Dur</th>
<th>PG Dur</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retrieval ability</td>
<td>0.095</td>
<td>-0.233</td>
<td>0.031</td>
<td>-0.008</td>
</tr>
<tr>
<td>Crystallized intelligence</td>
<td>-0.443</td>
<td>-0.569*</td>
<td>-0.159</td>
<td>-0.165</td>
</tr>
<tr>
<td>Fluid intelligence</td>
<td>-0.055</td>
<td>-0.144</td>
<td>0.077</td>
<td>-0.033</td>
</tr>
<tr>
<td>Creativity</td>
<td>-0.189</td>
<td>0.006</td>
<td>-0.197</td>
<td>-0.092</td>
</tr>
</tbody>
</table>

* p < .05

Note. Amp = peak amplitude; Dur = duration; LP = literal phrases; CM = conventional metaphors; NM = novel metaphors; PG = participant generated
Figure 1. ERP at Cz for all phrase types averaged across testing sessions (except PG which were only measured on Day 3).
Figure 2. Electrical signal (microvolts) at the N400 peak by phrase type and testing session.

Error bars represent standard deviations.
Figure 3. Duration of N400 peak (ms) by phrase type averaged across testing sessions (with the exception of participant-generated metaphors, which were only shown during testing session 3). Error bars represent standard deviations.
Figure 4. ERP at Cz comparing participant-generated metaphors to other metaphor types by day.
Appendix A

Examples of ERP Stimuli

<table>
<thead>
<tr>
<th>Literal Phrase</th>
<th>Conventional Metaphor</th>
<th>Novel Metaphor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Revolution is war</td>
<td>Argument is war</td>
<td>Addiction is war</td>
</tr>
<tr>
<td>Daisies are flowers</td>
<td>Friends are flowers</td>
<td>Smiles are flowers</td>
</tr>
<tr>
<td>Cobras are snakes</td>
<td>Lawyers are snakes</td>
<td>Torpedoes are snakes</td>
</tr>
<tr>
<td>Tango is a dance</td>
<td>Life is a dance</td>
<td>Traffic is a dance</td>
</tr>
<tr>
<td>A loofa is a sponge</td>
<td>Memory is a sponge</td>
<td>A tumor is a sponge</td>
</tr>
</tbody>
</table>

Examples of Verbal Cognitive Abilities Measurements

1) Creativity Measure: Remote Associates Test (RAT)

Look at the three words and find a fourth word that is related to all three.

TOOTH     POTATO     HEART     ____________
HOUND     PRESSURE   SHOT  ____________
MARSHAL    CHILD     PIANO     ____________
SENSE     COURTESY   PLACE   ____________

2) Crystallized intelligence

6. visor     7. hapless     8. chagrin
A) mask       A) accidental  a) anger
B) inspector  B) gloomy     B) revenge
C) clamp      C) unfortunate C) mortification
D) hawk       D) noiseless  D) disgust
E) snake      E) unknown    E) disgrace

3) Fluid intelligence: letter sets task

In the five groups of letters below, four of the groups have something in common. Mark the group which is not like the others.

1. VBCDF     WQRSH     XIKJF     YLMNF     ZABCF
2. ELMNE     FABCHE     GOPQG    HZTMH     IVWXI
3. LGBGM     TBKAK     ZDNEN     FRLLL   VPSPW
4. XVTQZ     BFDHJ     HDFBJ     JBFDH   HJFBD
5. BCAHV     BCOHV     DZELM     DKILM   DTCLM
4) Retrieval ability: verbal fluency test from the Cognitive Telephone Screening Instrument (COGTEL)

Please try to name as many words as possible that start with the letter K during 1 minute. You should not repeat any words and you should not say any names, for instance Klaus is not valid.

Scoring:
Number of named words ___
Number of names ___
Number of repeated words ___