Semantic Distance Modulates the N400 Event-Related Potential in Verbal Analogical Reasoning

Matthew J. Kmiecik (mkmiecik@luc.edu)

Department of Psychology Loyola University Chicago 1032 W. Sheridan Rd. Chicago, IL 60626 USA

Robert G. Morrison (rmorrison@luc.edu)

Department of Psychology, Neuroscience Institute Loyola University Chicago 1032 W. Sheridan Rd. Chicago, IL 60626 USA

Abstract

Computational accounts have traditionally focused on mapping between structured representations as fundamental to analogical processing. However, a recent connectionist model has been used to argue that structured representations may not be necessary to solve verbal analogies. Green and colleagues (2010) have shown that brain areas associated with analogical mapping become more engaged as semantic distance increases between verbal analogy source and targets. Herein, we had participants verify verbal analogies characterized for semantic distance while we monitored their brain waves using EEG. Our results suggest that the semantic distance between the source and target of a verbal analogy does influence early semantic processing as reflected in the N400 Event-Related Potential. However, successfully differentiating valid and invalid verbal analogies engages areas of prefrontal cortex widely associated with inhibitory processing and the integration of abstract relations in working memory. Thus, it appears that traditional semantic priming alone is likely insufficient to explain the full extent of analogical processing.

Keywords: analogy, semantic distance, EEG, N400

Introduction

Analogical reasoning is fundamental to the way that humans learn and reason in day-to-day life.. Likewise, analogies have long been considered to be a core component of analytic intelligence (Spearman, 1927) and of great importance in learning and discovery (Holyoak & Thagard, 1995). For nearly a century, researchers in cognitive science have developed theories and computational models to offer potential mechanisms for analogical processing (French, 2002). More recently patient-based (Morrison et al., 2004; Krawczyk et al., 2008) and functional neuroimaging studies (e.g., Bunge et al. 2005; Bunge et al. 2009; Green et al. 2010; Krawczyk et al., 2010; Volle et al., 2010; Watson & Chatterjee, 2012) have begun to identify a network of brain areas, particularly the prefrontal cortex (PFC), essential for analogical reasoning. Four-term verbal analogies have long been used as both a standard measure of intelligence and vocabulary knowledge. According to traditional accounts of analogical processing, to solve this type of problem the reasoner needs to (1) retrieve word meanings from semantic memory, (2) bind words into explicit abstract relations in working memory, and (3) perform a mapping in working memory between corresponding sets of words in the source and target. For instance, to verify the analogy:

animal : zoo :: person : house

participants may (1) retrieve the meanings of the words animal, zoo, person, and house, (2) bind *housed* (animal), *lived-in* (zoo), *housed* (person), and *lived-in* (house) (3) and then map *lives-in* (animal, zoo) to *lives-in* (person, house) specifically discovering that animal analogically maps to person and zoo maps to house. Several researchers have used this type of approach as embodied in the LISA model (Hummel & Holyoak, 1997; 2003) to account for patterns of verbal analogy performance (Morrison et al., 2004; Michael Vendetti & Knowlton & Holyoak, 2012).

In contrast, recent connectionist models of analogy (Leech, Mareschal & Cooper, 2008) have proposed that four-term verbal analogies may be solved without the use of structured relations via a mechanism utilizing guided pattern completion in semantic memory. Contrary to previous accounts of analogical priming (Spellman, Holyoak, & Morrison, 2001), Leech and colleagues argue that this mechanism of analogy could occur automatically without the use of explicitly represented relations and analogical mapping.

In addition to many experimental studies (see Holyoak & Hummel, 2008) the former traditional relationally explicit approach is supported by findings showing that solving verbal analogies engages anterior regions of the PFC (Bunge et al., 2005; Green et al., 2010) frequently associated with processing abstract information (e.g., Christoff et al., 2009; Nikitin & Morrison, 2011; Volle et



Figure 1: (a) Trial timeline. Participants were instructed to think of how the A:B were related and then decide whether the C:D pair was related in the same way. Calculated ERPs were time-locked to presentation of the C:D pair. (b) C:D pairs were used for all four conditions across four blocks of trials. Valid and invalid problems were matched for semantic distance using LSA for both near and far conditions.

al., 2010) and explicit relational integration (e.g., Bunge et al., 2009; Nikitin & Morrison, 2011; Volle et al., 2010; Watson & Chatterjee, 2012). However, a recent set of functional magnetic resonance imaging (fMRI) studies suggest that verbal analogical reasoning may exist on a continuum between the two approaches depending on the nature of the analogies. Green et al. (2010) developed a problem set of four-term verbal analogies that varied in the semantic distance between the source and target as measured using Latent Semantic Analysis (LSA, Landauer, Foltz, & Laham, 1998). Green and colleagues found that the anterior regions of PFC frequently associated with relational integration and/or abstract information processing were engaged to a greater extent when the source and target domains of the analogy were more distant ("far" analogies). This result suggests that "near" analogies may employ processing less dependent on structured representations.

To further explore this distinction we employed Green and colleagues' method of differentiating near and far analogies to develop a large set of verbal analogy problems for use with scalp electroencephalography (EEG). Researchers interested in the use of semantic memory during language processing have frequently used EEG analyzed with event-related potentials (ERPs) to investigate the time-course of semantic processing. Specifically, the N400 is a negative ERP component that typically peaks around 400ms after presentation of the stimulus. The N400 increases in negativity as a stimulus (usually a single word) becomes more incongruous from its context (Kutas & Federmeier, 2011). The N400 effect was first documented in sentence processing. For example, the italicized word in the sentence, "The cat will bake the food" will elicit a more negative N400 relative to, "The cat will eat the food" (Kutas & Hillyard, 1980). Many studies of language processing and semantic processing have shown the N400 to be sensitive to contextual semantic meaning (Kutas & Federmeier, 2011). Subsequent work has shown that the N400 effect is elicited in response to conceptual incongruities in other domains. For example, incorrect answers to simple symbolic (e.g., "4 x 4 = $21^{"}$) and verbal (e.g., "Twelve plus three equals sixteen.") arithmetic problems elicit an N400 effect (e.g., Niedeggen & Rosler, 1999).

Importantly, this type of automatic semantic congruity processing as measured by ERP methodology occurs earlier in the time course of processing than structured comparisons such as syntactic processing or analogical mapping. For instance, a positive ERP component typically peaking around 600ms after stimulus presentation (the P600) is sensitive to violations of syntax within a sentence (e.g., "The student will *studying* the lecture the professor gave on tuesday."; Osterhout & Mobley, 1995). Likewise the P600 is also sensitive to violations in structure of music (Patel et al., 1998). Likewise, Nikitin and Morrison (2011) found that an ERP component linked to the comparison of relational structures during visual analogical reasoning began approximately 500 to 600ms post stimulus presentation, once again after the N400.

To further explore the influence of semantic distance on analogical reasoning, we recorded EEG while participants solved sequentially presented four-term verbal analogy problems (e.g. A:B::CD; see Figure 1) varying in the semantic distance between the source (A:B) and target (C:D) word pairs. Semantic distances between the first and second word pairs were split into near (semantically similar) and far (semantically less similar) analogies. We hypothesized that near analogies would be more likely than far analogies to be solved via automatic semantic priming and thus the N400 ERP would be less negative for near than far analogies.

Method

Participants

Seventeen Loyola University Chicago undergraduate students (M=21.4 years old) participated in the experiment. Participants gave informed consent to take part in the study and were paid according to procedures approved by the Loyola University Chicago Institutional Review Board. One participant was excluded from the analysis due to poor EEG recording quality.

a)

Materials

Four-term verbal analogy problems were constructed from pairs of words representing one of five possible relations: *kept in* (e.g. animal:zoo), *kind of* (e.g. aluminum:metal), *made of* (e.g. candle:wax), *used to* (e.g. train:travel), and *works for* (e.g. curator:museum). To ensure that word pairs were representative of the claimed relation we had an earlier group of 10 participants perform a relation naming task with a candidate list of word pairs. In the present study we only included word pairs in which participants could quickly name the stated relation from the five possibilities. Word pairs with identical relations were paired to form valid analogies and pairs representing different relations were paired to form invalid analogies.

Following the methods of Green and colleagues (2010) we further divided valid and invalid analogies based on 4term semantic distance using Latent Semantic Analysis (LSA) into either near or far analogies. LSA performs complex algorithms on large corpora of text (semantic spaces) to produce semantic similarity ratings for pairs of words (Landauer, Foltz & Laham, 1998). The Matrix Comparison feature in LSA allows users to enter a list of nterms or word pairs and produce similarity ratings between all terms or pairs of words $(n \times n)$ entered in the list. Similarity ratings within the source (A:B) and target (C:D) of each analogy problem, as well as similarity ratings between the source and target, were acquired. The source word pairs and overall analogies were characterized for both near and far semantic distance using the obtained similarity ratings from LSA.

Two counterbalanced versions of 360 unique problems were created with 90 of each type of trial. For each version every C:D word pair was used in all four conditions. To minimize the chance of confounding the N400 due to repetition effects we divided each version into quarters so that CD word pairs could be separated in time, one in each quarter. Both valid and invalid problems differed with respect to semantic distance in the same way, and this was consistent across the quarters of the experiment. Importantly, problems did not differ with respect to mean word length or word frequency as measured using HAL (Burgess & Lund, 1997; Balota et al., 2007) across problem types.

EEG Recording Procedure

EEG was recorded from each participant using a 72-channel Biosemi Active2 EEG system. 64 electrodes were located at equidistant locations in a nylon cap. To expand the coverage of EEG monitoring we placed two electrodes on the inferior edge of the orbit of each eye. Raw EEG was re-referenced to an average of the two mastoid electrodes and then highpass filtered at 0.01 Hz. The signal was then band-stop filtered from 59 to 61 Hz to remove any AC electrical contamination. EEG signal was corrected for ocular artifacts using a spatial PCA filter, a method available in EMSE (Source Signal Imaging, San Diego CA). Signal was further cleaned via a $\pm 100\mu$ V rejection criterion. Included participants had fewer than 15% of trials rejected due to EEG artifacts. A 20Hz low-pass filter was applied to ERPs for visualization only.

Procedure

After being prepared for the EEG recording, participants sat in a quiet room equipped with a 21-inch CRT monitor and an electronic response box. Participants sat 100cm from the monitor. Stimulus width was adjusted to 4 degrees of visual angle. The task was run and data were collected using e-Prime 2.0 (Psychology Software Tools, Sharpsburg, PA).

After task instructions the participant performed 20 practice trials with feedback. Each trial began with a randomly jittered fixation screen that lasted 500 to 1000ms (See Figure 1a). The first word pair was presented at the center of the screen for 3.3s. Participants were instructed to think of how the pair of words was related. Following an equal sign presented for 750ms, a second pair of words was presented for 3.5s during which participants were decide whether the two pairs of words were related in the same way (i.e., formed a valid analogy). Participants indicated their choice by pressing one of two buttons with two fingers from their right hand. The entire experiment consisted of 360 trials divided into twelve blocks separated by 20s rests.

Results

Behavioral Results

Participants were more accurate in judging near than far analogies (see Figure 2a; F(1,15) = 28.6, p < .001, $\eta_p^2 = .66$). There was no difference in accuracy with respect to validity (F(1,15) = .2, ns, $\eta_p^2 = .01$); however, there was a significant interaction between semantic distance and validity (F(1,15) = 8.5, p = .01, $\eta_p^2 = .36$). Further contrasts suggested that this interaction was driven by participants being more accurate for valid near than valid far problems (F(1,15) = 24.9, p < .001, $\eta_p^2 = .62$). Participants were also faster in judging near than far

Participants were also faster in judging near than far analogies (see Figure 2b; F(1,15) = 12.1, p = .003, $\eta_p^2 = .45$), faster in judging valid compared to invalid problems (F(1,15) = 9.7, p = .007, $\eta_p^2 = .39$), and the two factors also interacted (F(1,15) = 15.9, p = .001, $\eta_p^2 = .51$). The interaction was driven by participants being faster for valid near than valid far problems (F(1,15) = 22.6, p < .001, $\eta_p^2 = .60$) and faster for valid than invalid near problems (F(1,15) = 18.5, p = .001, $\eta_p^2 = .55$).



Figure 2: Verbal analogy (a) accuracy and (b) RT.

EEG Results

We calculated grandaverage ERPs for correct trials for each of the four conditions (see Figure 3a). Initially we divided the first 1400ms of processing into seven 200ms epochs and performed a 2 (near vs. far) x 2 (valid vs. invalid) x 7 (Time) repeated measures ANOVA on mean amplitudes from a central electrode (i.e., Cz) frequently used in N400 analyses. There were reliable main effects of semantic distance $(F(1,15) = 20., p < .001, \eta_p^2 = .57)$, and time $(F(6,90) = 11.8., p < .001, \eta_p^2 = .44)$, and a trend towards a difference based on the validity of the analogy (F(1,15) =3.5., p = .08, $\eta_p^2 = .19$). Importantly there was a three way interaction of type, validity and time (*F*(6,90) = 5.1., *p* < .001, $\eta_p^2 = .26$). As can be seen in Figure 3a far problems have more negative ERPs beginning around the N400; however, near invalid problems later join far valid and invalid problems as being more negative than near valid problems. While topographies based on valid/invalid subtractions (see Figure 3b) are broad they tend to suggest that differences in the near valid vs. invalid conditions move from more central to more right frontal distributions.

Second, we focused on the N400 and calculated mean amplitude for an early 300-500ms time window typically used for analyzing the N400 in studies of semantic priming. We ran a 2 (near vs. far) x 2 (valid vs. invalid) repeated measures ANOVA. The N400 as measured in this time window was more negative for far than near problems regardless of problem validity ($F(1,15) = 19.5., p < .001, \eta_p^2 = .57$), with no interaction ($F(1,15) = 1.5, p = .24, \eta_p^2 = .09$).

To compare our results to those of Green and colleagues (2010) and to attempt to understand the time course of neural activity with respect to semantic distance and topography, we conducted an additional analysis on just correct near and far valid problems (see Figure 4). In this analysis we focused on the early 300-500ms time window previously mentioned and a later 900-1100ms time window closer to the response. Nikitin and Morrison (2011) have previously shown this later time window to be associated with analogical mapping in a visual analogy task. Adapting

the methods of McCarthy and Woods (1985) we normalized the subtraction of near and far mean amplitudes for each time window. A 2 (near vs. far) x 2 (early vs. late) repeated measures ANOVA demonstrated that the normalized near/far subtraction reversed from initially being greater in central areas to later being greater in frontal areas yielding a reliable interaction (see Figure 4b; F(1,15) = 9.6, p = .007, $\eta_p^2 = .39$). Normalized subtractions showed an increase in frontal channels over time (F(1,15) = 5.9, p = .02, $\eta_p^2 = .28$), while central channels showed a trend towards a decrease (F(1,15) = 3.6, p = .08, $\eta_p^2 = .19$). Thus, we believe that Green and colleagues (2010) result that frontopolar areas were more active for far than near analogies may be driven by later processing likely reflective of the greater reliance on analogical mapping while solving far analogies.

Discussion

As hypothesized, we found that the semantic distance between source and target word pairs in verbal analogy problems modulated the mean amplitude of the N400 ERP with near analogies eliciting less negative N400s compared to far analogies. The N400 ERP is sensitive to word repetition, semantic integration, and semantic expectancy effects (Kutas & Federmeier, 2011). Controlling word repetition effects by utilizing identical second word pairs across all four conditions ensured semantic integration processes were isolated when analyzing N400 modulation. A more negative N400 for far analogies can be explained by the 'knowledge integration effort' view, which suggests negativity in N400 amplitude is directly proportional to the integration effort required to extract lexical representations for each target (Holcomb, 1993). Increases in semantic integration effort in far analogies were reflected in more negative N400 mean amplitude as semantic distance between source and target analogs increased.

The knowledge integration effort view also explains the less negative N400 mean amplitude observed in near analogies. As semantic distance between source and target analogs decreased, less semantic integration effort was



Figure 3: a) Grandaverage stimulus-locked ERPs (electrode Cz) for correct Valid and Invalid, near and far analogies. b) Topographic maps of valid/invalid subtractions for near and far analogies across the time course.





Figure 4: a) Stimulus-locked ERPs for correct valid near and far conditions. Grandaverage ERPs for each condition were calculated from clusters of frontal and central electrodes indicated above. b) Topographic maps of near/far subtractions for the early and late time windows.

required to derive lexical representations. In other words, deriving the semantic information of the first word pair and determining the source analog relation facilitated access to the second word pair, resulting in a less negative N400 in near analogies. A previous study demonstrated that analogous source pairs facilitated lexical access to target words when participants were instructed to attend to and use relational information (Spellman, Holyoak & Morrison, 2001). Since participants were solving analogies and attending to the relations between words, analogical priming

may have facilitated semantic integration particularly in the near analogy condition.

However, while semantic priming may be sufficient to explain near valid analogies it may not be sufficient to reject false analogies or perform far analogies. Beginning at the N400, valid and invalid ERPs for near analogies diverge. Closer inspection of the topographies (see Figure 4) suggests engagement of areas of the brain traditionally associated with inhibitory processing during analogy (Cho et al. 2010; Watson & Chatterjee, 2012). In fact, Morrison and colleagues (2004) have previously demonstrated that frontal patients have great difficulty rejecting lures in twochoice verbal analogy problems where semantic congruity for the false item is greater than for the true item.

Likewise, while far analogies do show a more negative N400 than near analogies, suggesting that automatic semantic processing is indexing semantic distance in analogy, there is no difference in the N400 between far valid and invalid analogies demonstrating that semantics alone are insufficient for complete analogical processing. In fact, like invalid near condition analogies, far analogies engage prefrontal cortex to a greater extent, consistent with findings by Green and colleagues (2008).

Thus, our findings suggest approaches relying solely on tranditional semantic priming, such as recent connectionist approaches (e.g., Leech, Mareschal & Cooper, 2008), have limited applicability when the distance between the source and targets of analogies increases, or when the reasoner must choose between alternative analogues where semantics alone do not indicate the more relationally similar match.

Acknowledgements

The authors would like to thank Ryan Brisson and Krishna Bharani for technical support. The Provost and the Mulchay Scholars Program of Loyola University Chicago (MJK, RGM), the American Federation of Aging Research/Rosalinde and Arthur Gilbert Foundation (RGM), and the Illinois Department of Public Health (RGM) provided support.

References

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., . . . Treiman, R. (2007). The english lexicon project. *Behavior Research Methods*, *39*(3), 445-459.

Bunge, S. A., Helskog, E. H., & Wendelken, C. (2009).

- Left, but not right, rostrolateral prefrontal cortex meets a stringent test of the relational integration hypothesis. *NeuroImage*, 46(1), 338-342.
- Bunge, S. A., Wendelken, C., Badre, D., & Wagner, A. D. (2005). Analogical reasoning and prefrontal cortex: Evidence for separable retrieval and integration mechanisms. *Cerebral Cortex (New York, N.Y.: 1991)*, 15(3), 239-249.
- Burgess, C., & Lund, K. (1997). Modeling parsing constraints with high-dimensional context space. *Language and Cognitive Processes*, 12, 177-210.

- Cho, S., Moody, T. D., Fernandino, L., Mumford, J. A., Poldrack, R. A., Cannon, T. D., Knowlton, B. J., & Holyoak, K. J. (2010). Common and dissociable prefrontal loci associated with component mechanisms of analogical reasoning. *Cerebral Cortex*, 20, 524-533.
- Christoff, K., Keramatian, K., Gordon, A. M., Smith, R., & Madler, B. (2009). Prefrontal organization of cognitive control according to levels of abstraction. *Brain Research*, *1286*, 94-105.
- French, R. M. (2002). The computational modeling of analogy-making. *Trends in Cognitive Sciences*, 6(5), 200-205.
- Green, A. E., Kraemer, D. J. M., Fugelsang, J. A., Gray, J. R., & Dunbar, K. N. (2010). Connecting long distance: Semantic distance in analogical reasoning modulates frontopolar cortex activity. *Cerebral Cortex*, 20(1), 70-76.
- Holcomb, P. J., & Anderson, J. E. (1993). Cross-modal semantic priming: A time-course analysis using event-related brain potentials. *Language and Cognitive Processes*, 8(4), 379-411.
- Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA, US: The MIT Press.
- Holyoak, K. J.,& Hummel, J. E. (2008). No way to start a space program. Commentary on R. Leech, D. Mareschal and R. P. Cooper, Analogy as relational priming: A developmental and computational perspective on the origins of a complex skill, *Behavioral and Brain Sciences*, 31, 388-389.
- Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review*, *104*(3), 427-466.
- Hummel, J. E., & Holyoak, K. J. (2003). A symbolicconnectionist theory of relational inference and generalization. *Psychological Review*, 110(2), 220-264.
- Krawczyk, D. C., Morrison, R. G., Viskontas, I., Holyoak, K. J., Chow, T. W., Mendez, M. F., . . . Knowlton, B. J. (2008). Distraction during relational reasoning: The role of prefrontal cortex in interference control. *Neuropsychologia*, 46(7), 2020-2032.
- Kutas, M. & Federmeier, K.D. (2011) Thirty years and counting: Finding meaning in the N400 component of the Event-Related Brain potential (ERP), *Annual Review of Psychology* Vol. 62: 621-647.
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207(4427), 203-205.
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, 25(2-3), 259-284.
- Leech, R., Mareschal, D., & Cooper, R. P. (2008). Analogy as relational priming: A developmental and computational perspective on the origins of a complex cognitive skill. *Behavioral and Brain Sciences*, 31(4), 375-378.
- McCarthy, G., & Wood, C. C. ~1985!. Scalp distributions of event-related potentials: An ambiguity associated with

analysis of variance models. *Electroencephalography and Clinical Neurophysiology, 62,* 203–208.

- Morrison, R. G., Krawczyk, D. C., Holyoak, K. J., Hummel, J. E., Chow, T. W., Miller, B. L., & Knowlton, B. J. (2004). A neurocomputational model of analogical reasoning and its breakdown in frontotemporal lobar degeneration. *Journal of Cognitive Neuroscience*, 16(2), 260-271.
- Morrison, R.G., Holyoak, K.J., & Truong, B. (2001). Working memory modularity in analogical reasoning. *Proceedings of the Twenty-fourth Annual Conference of the Cognitive Science Society* (pp. 663-668). Mahwah, NJ: Erlbaum.
- Niedeggen, M., Rosler, F., & Jost, K. (1999). Processing of incongruous mental calculation problems: Evidence for an arithmetic N400 effect. *Psychophysiology*, *36*(3), 307-324.
- Nikitin, S., & Morrison, R. G. (2011). Analogical reasoning in human prefrontal cortex: an event-related potential approach. Cognitive Neuroscience Annual Meeting, San Francisco, CA.
- Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory* and Language, 34(6), 739-773.
- Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience*, 10(6), 717-733.
- Spearman, C. (1927). *The abilities of man; their nature and measurement*. New York: Macmillan Co.
- Spellman, B. A., Holyoak, K. J., & Morrison, R. G. (2001). Analogical priming via semantic relations. *Memory & Cognition*, 29(3), 383-393.
- Vendetti, M., Knowlton, B. J., & Holyoak, K. J. (2012). The impact of semantic distance and induced stress on analogical reasoning. *Cognitive, Affective, and Behavioral Neuroscience*, 12, 804-812.
- Volle, E., Gilbert, S. J., Benoit, R. G., & Burgess, P. W. (2010). Specialization of the rostral prefrontal cortex for distinct analogy processes. *Cerebral Cortex*, 20(11), 2647-2659.
- Watson, C.E. & Chatterjee, A. (2012). A bilateral frontoparietal network underlies visuospatial analogical reasoning. *NeuroImage*, 59: 2831-2838