Published in final edited form as:

Psychophysiology. 2017 June ; 54(6): 833–847. doi:10.1111/psyp.12840.

Flexible conceptual combination: Electrophysiological correlates and consequences for associative memory

Heather D. Lucas^a, Ryan J. Hubbard^{a,b}, and Kara D. Federmeier^{a,b}

^aBeckman Institute for Advanced Science and Technology, University of Illinois, Urbana, Illinois, USA

^bDepartment of Psychology, University of Illinois, Champaign, Illinois, USA

Abstract

When meaningful stimuli such as words are encountered in groups or pairs (e.g., “elephant-ferry”), they can be processed either separately or as an integrated concept (“an elephant ferry”). Prior research suggests that memory for integrated associations is supported by different mechanisms than is memory for nonintegrated associations. However, little is known about the neurocognitive mechanisms that support the integration of novel stimulus pairs. We recorded ERPs while participants memorized sequentially presented, unrelated noun pairs using a strategy that either did or did not involve attempting to construct coherent definitions. We varied the concreteness of the first noun in each pair to examine whether conceptual combination instructions would induce compositional concreteness effects, or differences in ERPs evoked by the second noun as a function of the concreteness of the first noun. We found that the conceptual combination task, but not the noncombinatory encoding task, produced compositional concreteness effects on a late frontal negativity previously linked to visual imagery. Moreover, word pairs studied under conceptual combination instructions showed evidence of more unitized or holistic memory representations on associative recognition and free recall tests. Finally, item analyses indicated that (a) items with higher normed imageability ratings were rated by participants as easier to conceptually combine, and (b) in the conceptual combination task, ease-of-combination ratings mediated an indirect relationship between imageability and subsequent associative memory. These data are suggestive of a role of compositional imagery in the online formation of novel concepts via conceptual combination.

Descriptors

Conceptual combination; Associative memory; Unitization; ERPs; Imagery

To navigate life effectively, we must constantly learn new associations among items, concepts, and other elements of experience. In some cases, these associations are between

Address correspondence to: Heather D. Lucas, Ph.D., Beckman Institute for Advanced Science and Technology, 405 N Mathews Ave., Urbana, IL 61801, USA. hdlucas@gmail.com.

Supporting Information

Additional supporting information may be found in the online version of this article:

Appendix S1. Study-phase ratings, associative recognition performance, and the relationship between study-phase ratings and associative memory for Experiment 2.

relatively arbitrary and independent entities, such as an event and the date on which it will take place. By contrast, another form of associative processing— which is here referred to as *conceptual combination* (e.g., Del Prato & Pykkänen, 2014; Gagné & Spalding, 2013; Graves, Binder, Desai, Conant, & Seidenberg, 2010), and has elsewhere been termed relational integration (Estes & Jones, 2009) or conceptual fusion (Kounios, Smith, & Yang, 2001)—can occur when associated stimuli are perceived to modify or alter the properties of one another, leading to a combined interpretation that is more than the sum of its constituents. For example, while the phrase “house by a tree” suggests two relatively independent entities associated by location, the phrase “house in a tree” invokes a relationship (e.g., that of a “tree house”) that is integrative rather than additive.

Research on episodic memory suggests that associations that are initially processed via conceptual combination can be represented in memory in a manner that is more unitized or holistic than are nonintegrative associations¹ (Haskins, Yonelinas, Quamme, & Ranganath, 2008; Parks & Yonelinas, 2015; Quamme, Yonelinas, & Norman, 2007). In these studies, higher levels of unitization² have been inferred when the neural and/or behavioral properties observed during associative retrieval resemble those typically linked to the retrieval of individual items. For example, numerous functional neuroimaging (e.g., Hannula & Ranganath, 2008) and neuropsychological studies (e.g., Konkelt, Warren, Duff, Tranel, & Cohen, 2008) indicate that the hippocampus is critically involved in the retrieval of arbitrary (nonunitized) associations. By contrast, the adjacent rhinal cortex—particularly the perirhinal cortex, which has often been linked to single-item memory (e.g., Davachi, 2006)—has been implicated in the retrieval of associations that were subject to conceptual combination (Haskins et al., 2008; Quamme et al., 2007). Moreover, Parks and Yonelinas (2015) showed that (a) word pairs encoded via conceptual combination subsequently promoted familiarity-based associative recognition (a form of recognition that is thought to primarily support item memory) more so than word pairs subject to other types of deep encoding, and (b) conceptual combination-based encoding selectively enhanced associative memory (e.g., memory for pairs) without affecting memory for individual items.³ Importantly, in several studies (Bader, Mecklinger, Hoppstädter, & Meyer, 2010; Haskins et al., 2008; Parks & Yonelinas, 2015; Quamme et al., 2007), such evidence for higher levels of unitization was found even for arbitrary and unrelated noun pairs (e.g., “steam token”) following a single exposure in which these pairs were introduced to participants as novel compound words (“a coin used to purchase a spa’s steam bath”).

The notion that conceptual combination can allow neurocognitive mechanisms typically associated with item memory to support memory for associations has potential clinical implications. Relative to item memory, associative memory is disproportionately

¹Note that the term *unitization* can refer to the formation of a holistic memory representation for any configuration of stimuli, including relatively nonverbalizable stimuli (e.g., Rubin, Chesney, Cohen, & Gonsalves, 2013). Thus, conceptual combination is only one form of processing that can lead to unitization.

²The phrase *levels of unitization* is used here to reflect the notion that the unitization process may be continuous, rather than all or none, and that no established benchmark separates associations that can versus cannot be said to be unitized (Parks & Yonelinas, 2015). It is thus more accurate to refer to certain stimuli as being “higher” or “lower” in unitization relative to other stimuli than it is to discuss unitization in binary or absolute terms.

³The latter of these findings distinguishes the unitizing effects of conceptual combination from those of other beneficial encoding strategies such as depth-of-processing manipulations, which would be expected to increase both item and associative memory.

compromised by the aging process (e.g., Castel & Craik, 2003) and many neurological and psychiatric diseases (e.g., Monti et al., 2013; Williams et al., 2010). Thus, techniques similar to conceptual combination may help to remediate associative memory deficits in these populations (e.g., Ahmad, Fernandes, & Hockley, 2014; Bastin et al., 2013). However, the ability to make informed decisions about when and for whom conceptual combination is best suited as a memorization strategy requires a clear understanding of its underlying mechanisms, particularly as they apply to associations that are relatively arbitrary. Conceptual combination is also an important topic of study in its own right, as it constitutes a tool for efficiently organizing experiences with the world into novel, meaningful units. Here, we investigate the electrophysiological mechanisms of what we term *flexible* conceptual combination, or the semantic integration of novel word pairs with neither preexisting nor unambiguous interpretations.

Our focus on flexible conceptual combination contrasts with that of previous work, which has focused largely on what has been termed *fluent* conceptual combination, or the processing of stimuli that cohere in a natural and unambiguous manner (e.g., “lake house”; Bemis & Pykkänen, 2013; Graves et al., 2010; Molinaro, Carreiras, & Duñabeitia, 2012). Although some studies have examined conceptual combination for novel word pairs (as in the above steam token example) these studies have focused primarily on ways in which retrieval processes are affected by prior conceptual combination rather than on the conceptual combination process itself (but see Haskins et al., 2008). Moreover, participants in each of these studies were provided with experimenter-generated definitions for each stimulus pair. This design choice may have minimized the need for the participants to engage in certain aspects of flexible conceptual combination, such as the broad activation of knowledge pertaining to the individual words and the identification of a plausible relational structure through which to integrate them.

Of particular interest in the present study is the extent to which the cognitive properties and temporal dynamics of flexible conceptual combination resemble those that have previously been found in studies of fluent conceptual combination. Previous studies have examined fluent conceptual combination by contrasting patterns of electrophysiological activity elicited by modifier-head pairs with versus without clear combined meanings (e.g., “lonely monster” vs. “geographic monster”), under the assumption that combinatorial processing mechanisms will be preferentially engaged in the former case. For example, in an ERP study using Spanish nounadjective modification rules (Molinaro et al., 2012), N400 potentials—an electrophysiological index of the processing demands associated with semantic access (Kutas & Federmeier, 2011)—were reduced in amplitude for adjectives that could be coherently linked to their preceding head nouns, regardless of whether the adjectives denoted features that were high typicality (“horrible monster”) or low typicality (“lovely monster”). This finding converges with behavioral evidence from lexical decision tasks (Estes & Jones, 2009; Gagné, Spalding, & Ji, 2005) to suggest that when a word is identified as the leading word of a compound, information about its modifying roles (for modifiers) or modifiable properties (for heads) becomes active and facilitates the semantic processing of likely companion words. However, these data do not speak to the mechanisms by which more remote or unusual modifier-head relationships are processed. Again taking steam token as an example, the noun *steam* tends to modify other nouns via the *has* relations (as in

“steam room”) or the *uses* relation (as in “steam engine”), but rarely the *buys* relation. Accordingly, it is plausible that neurocognitive processes subsequent to those indexed by the N400 enable the representational flexibility needed to arrive at a “token buys steam” interpretation.

Huang, Lee, and Federmeier (2010) developed a paradigm for studying conceptual combination that does not require the use of low-coherence stimuli as a baseline, and can therefore be adapted to specifically study the integration of words with low preexperimental coherence. These authors investigated ways in which ERPs recorded in response to the same lexical head differ according to the concreteness of the word that modifies it. Extensive prior research has documented multiple ERP subcomponents that are influenced by word-level concreteness, including N400 potentials as well as a late frontal negativity sometimes referred to as N700 (Gullick, Mitra, & Coch, 2013; Holcomb, Kounios, Anderson, & West, 1999; Welcome, Paivio, McRae, & Joanisse, 2011; West & Holcomb, 2000). Importantly, Huang et al. (2010) found that these ERP concreteness effects were also present in healthy young adults at the compositional level, or on the same head noun as a function of whether the preceding modifier activated relatively concrete versus abstract information that could be applied to that noun. Specifically, reading nouns preceded by concrete modifiers (e.g., “green book”) led to reduced N400 and enhanced N700 potentials relative to reading the same nouns preceded by abstract modifiers (e.g., “engaging book”). These effects are compositional because they indicate that the meaning of the modifier has become integrated into that of the head lexical item, such that concreteness differences are present at the message level. Moreover—at least in the context of that study, in which participants performed a coherence-judgment task—these effects were contingent on preexperimental coherence. For example, neither ERP component differentiated “watery glove” from “loyal glove” at the compositional level.

Together, these data suggest that ERP compositional concreteness effects can provide a window into the temporal dynamics of conceptual combinatorial processing, including how these dynamics vary by task demands and stimulus factors. Interestingly, concreteness effects on the N700 seem to be functionally dissociable from those on the N400, with the former having been linked to the greater use of sensory imagery when processing concrete concepts. In prior work, N700 effects have been found to be greatest during tasks that emphasize visualization, and smaller or absent for lexical decision tasks or tasks that emphasize surface-level processing (Gullick et al., 2013; Welcome et al., 2011; West & Holcomb, 2000). Moreover, research using visual half-field methods (Huang et al., 2010) demonstrates that, unlike N400 concreteness effects, N700 effects are maximal during left visual field/right hemisphere presentation, consistent with right hemisphere involvement in the use of sensory imagery during language processing (Paivio, 1991). As such, while N400 compositional concreteness effects may be relatively stimulus driven (e.g., resulting from the tendency of concrete modifiers to preactivate semantic features of likely head items), N700 effects may reflect imagistic processing that is more top down and thus able to contribute to conceptual combination for a wider range of stimulus pairs (e.g., Ishai, Ungerleider, & Haxby, 2000; Mechelli, Price, Friston, & Ishai, 2004).

A primary goal of the present study was to assess the mechanisms of flexible conceptual combination and its consequences for memory by recording ERPs as participants studied unrelated noun pairs under task instructions that either did or did not emphasize conceptual combination. In conceptual combination blocks, participants were asked to generate plausible compound definitions for each of the word pairs and provide trial-by-trial ratings of their success in doing so. By contrast, in frequency-comparison blocks, participants judged whether the concept denoted by the first word (W1) was one that they encountered more or less often relative to the concept denoted by the second word (W2). We chose frequency-comparison as a noncombinatory encoding task because it encourages meaning-based or “deep” processing of both of the items in each pair without encouraging their meanings to be combined. In addition, we adopted the methods used by Huang and colleagues of manipulating the concreteness of the W1 in each pair. Thus, for each participant within each task, an equal number of noun pairs were abstract-concrete (“idea salad”) and concrete-concrete (“road salad”).

Note that the conceptual combination task employed here differs from the coherence-judgment task used by Huang and colleagues in that the present task encourages participants to generate meanings for word pairs rather than merely judging the presence or absence of such meaning. For this reason, we expect to find certain compositional concreteness effects despite our use of low-coherence stimulus pairs. As previously mentioned, compositional concreteness effects on N400 likely emerged for coherent adjective-noun pairs in Huang et al. (2010) because concrete adjectives provided a more constraining semantic context that allowed participants to preemptively activate features of upcoming nouns. Because the W1s used in the present study were unrelated to the W2s, we did not expect to find N400 compositional concreteness effects. We did, however, predict that compositional concreteness effects would emerge on N700 potentials. In particular, in the conceptual combination task, we expected to observe more negative N700 amplitudes to W2s preceded by concrete relative to abstract W1s, perhaps reflecting a greater elicitation of imagery during the production and/or evaluation of combined meanings for concrete-concrete word pairs. Finally, we did not expect the same stimulus combinations to produce N700 compositional concreteness effects during the frequency-comparison task, since the W1 in this task should not influence the meaning elicited by the corresponding W2.

Given our interest in the relationship between flexible conceptual combination and memory outcomes, we examined word pair memory in multiple ways. First, we assessed associative memory accuracy for each block via an associative recognition test (Experiment 1). Note that, although unitization has been linked to better memory for word pairs (e.g., Parks & Yonelinas, 2015), improvements in associative memory do not, in and of themselves, make clear whether or not information has been unitized. To provide additional preliminary evidence of successful unitization, we performed a second behavioral experiment (Experiment 2). Here, we administered both study tasks to a separate group of participants and then tested their memory using an unstructured free recall test, in which participants were instructed to recall as many words as possible without regard to the order or pairings in which they had previously been presented. Previous research has found that task instructions intended to promote high levels of unitization lead to a selective improvement of memory for word pairs, without appreciably affecting memory for words presented in isolation (e.g.,

Bower, 1970; McGee & Rob, 1980; Parks & Yonelinas, 2015). Thus, we did not expect a difference in the overall number of words recalled between the two tasks. Rather, we reasoned that greater levels of unitization would manifest as an increase in within-pair conditional recall probabilities—or the extent to which recall of one word within a pair was contingent on recall of the other—in the conceptual combination relative to the frequency-comparison task. Such a pattern would be consistent with the notion that our self-directed combination task promoted unitization, similar to previous studies in which conceptual combination was experimenter directed (e.g., Bader et al., 2010; Haskins et al., 2008; Parks & Yonelinas, 2015; Quamme et al., 2007). Finally, we conducted item-level analyses to examine whether certain stimulus properties associated with either the W1 or the W2 systematically predicted conceptual combination success.

Experiment 1

Method

Participants—Twenty-four young adults from the University of Illinois and surrounding areas (18 female, mean age = 21, range = 18–24) participated in the experiment and were compensated \$10/hr. All were right-handed and native speakers of English.

Stimuli—The stimuli consisted of 144 noun pairs (72 abstract-concrete pairs and 72 concrete-concrete pairs) formed from 72 abstract nouns and 216 concrete nouns. All words were between four and eight letters in length. Abstract nouns had a mean concreteness rating of 280 (range = 232–349), and concrete nouns had a mean concreteness rating of 574 (range = 500–646), according to the Medical Research Council database described by Coltheart (2007). Concreteness ratings in this database range from 100–700.

To construct the word pairs, the 72 abstract nouns were first divided into four lists of 18 words each, matched for concreteness, imageability, familiarity, frequency, and word length. For each abstract noun, a “yoked” concrete noun was then chosen that was of comparable familiarity, frequency, and length to the abstract noun. These eight lists of 18 words served as the first words (W1s) of the 144 word pairs. Analyses of the final sets confirmed that each set of 18 abstract W1s differed significantly from its set of yoked concrete W1s in concreteness and imageability, but not in familiarity, frequency, or length.

The remaining 144 concrete words were divided into four lists of 36 words matched for concreteness, imageability, familiarity, frequency, and length. These words served as the second words (W2s) for all word pairs. Within each list of 36, two randomly chosen W2s were assigned to each pair of yoked W1s. This procedure allowed us to form two sets of word pairings, such that the frequency with which each W2 was preceded by a concrete versus an abstract W1 was counterbalanced across participants. For example, for every participant who received the pairs “grief-crew” and “pony-camera,” another participant saw “pony-crew” and griefcamera.” The frequency with which each of the four sets of words was presented in a conceptual combination block versus a frequency-comparison block was also counterbalanced across participants.

We manually inspected and adjusted noun pairings in order to eliminate any pairs that formed clear preexisting concepts (e.g., “flower store”). Moreover, free association data were available for 109 of the 144 W1s via the University of South Florida Free Association Norms database (Nelson, McEvoy, & Schreiber, 2004). For these words, we confirmed that there was negligible forward association strength from each W1 to its corresponding W2.

Procedures—The experiment consisted of four study-test blocks, two of which were conceptual combination blocks and two of which were frequency-comparison blocks. Each participant saw both blocks of each type consecutively to minimize task switching, and presentation order was counterbalanced. In each study phase, 36 word pairs were presented in a random order, bounded by one primacy buffer and one recency buffer. Within each block, half of these word pairs were abstract-concrete pairs, and the remaining half were concrete-concrete pairs. After a 30-s distractor task, all 36 word pairs were again presented in an associative recognition test. At test, 18 of the word pairs (nine abstract-concrete and nine concrete-concrete) were intact, meaning that they were paired together in the same order as they had been at study. The remaining pairs were re-paired, or paired with a different word from the same block. The frequency with which each word pair was presented as intact versus re-paired was counterbalanced across participants. Within each counterbalance, re-pairings were determined randomly for each participant, with the constraints that (a) W1 words remained first in each pair and were always re-paired with W2 words, and (b) words from abstract-concrete pairs were never repaired with words from concrete-concrete pairs, and vice versa. Each test block began with two practice buffer trials consisting of the primacy and recency buffers from the previous study block, which were randomly presented as either intact or re-paired.

Sample study trials for the conceptual combination and frequency-comparison task are depicted in Figure 1A,B, respectively. As shown, each study trial began with a 500-ms presentation of a W1, followed by an additional 1,000 ms of fixation, and then a 500-ms presentation of the corresponding W2. Next, a fixation cross appeared for 1,000 ms, followed by a 5,000-ms prompt indicating that the participant should respond. Participants were instructed to make their responses before the prompt disappeared.

During the conceptual combination blocks, each participant was instructed to rate on a 1–6 scale how well he or she could invent a definition for the compound phrase formed from the two words. Half of the participants were told to assign a rating of 1 to word pairs that they could not define at all (e.g., they had no idea what a compound phrase comprising the two words might mean), and to assign a rating of 6 to word pairs that were the easiest to clearly define. Buttons 2–5 reflected intermediate levels of difficulty, and were to be used when the participant had a more vague sense of what the corresponding phrase might mean. The mapping of buttons to levels of combination success was reversed for the remaining participants. Participants were asked to make these ratings on a relative basis and to attempt to use all six buttons approximately evenly. This instruction was given to help ensure that the ratings would provide information about which word pairs were relatively easier or harder to combine than others, regardless of each participant’s overall perception of task difficulty.

During the frequency-comparison blocks, participants were instructed to use a 1–6 scale to indicate which of the items denoted by the two words he or she encounters more frequently. For half of the participants, a rating of 1 indicated a much greater frequency of the first word, while a rating of 6 indicated a much greater frequency of the second word. Intermediate buttons were used for frequency differences that were less pronounced. These instructions were reversed for the remaining participants.

The timing and structure of the test trials were identical to those of the study trials (see Figure 1C for example), except that each participant was instructed to rate his or her confidence as to whether the preceding word pair was intact or re-paired on a scale of 1–6. For half of the participants, Button 1 was used to indicate a high level of confidence that the pair was intact, while Button 6 indicated a high level of confidence that the pair was rearranged. Button 2 and 5 were used to indicate medium levels of confidence, and Button 3 and 4 corresponded to low-confidence responses. This assignment was reversed for the remaining participants.

Prior to beginning the experiment, participants completed two practice study blocks, one with the conceptual combination instructions and the other with the frequency-comparison instructions. Practice study blocks contained three abstract-concrete pairs and three concrete-concrete pairs. During the practice conceptual combination block, participants were asked to verbally justify each response by providing the experimenter with a compound definition for the word pair. Whenever a participant was unable to come up with a compound definition, the experimenter offered 2–3 different examples of possible definitions to the participant to reinforce the instructions. Finally, participants completed a practice recognition test block containing intact and re-paired word pairs from each of the two study blocks.

ERPs were extracted from scalp EEG recordings from 26 Ag/AgCl electrodes spaced evenly over the head. Voltages were referenced online to a left mastoid electrode and rereferenced offline to averaged left and right mastoids. Eye movements and blinks were recorded from three additional electrodes below the center of the left eye and on the outer canthus of each eye. Electrode impedance was kept below 5 k Ω . Signals were recorded with a band-pass of .01–100 and sampled at a rate of 1000 Hz (BrainVision system). The mean amplitude of a 200-ms window prior to stimulus onset was subtracted to correct for baseline variability. Epochs containing electroocular or other artifacts were excluded from ERP analyses (mean = 9% of trials, range = 1–21%). Averaged ERP waveforms were band-pass filtered from 0.1–30 Hz prior to statistical analyses. Statistical comparisons were performed on amplitudes averaged over an anterior frontal electrode cluster (channels MiPf, LLPf, RLPf, LMPf, RMPf), a frontocentral electrode cluster (channels LDFr, RDFr, LMFr, RMFr, MiCe), and a parietal electrode cluster (MiPa, LMCE, RMCE, LDPa, RDPa). In these abbreviations, the first letter(s) denotes left (L), right (R), and midline electrodes (Mi). The second letter describes lateral (L), medial (M), and dorsal (D) locations. The final letters denote anteriority, as prefrontal (Pf), frontal (Fr), central (Ce), and parietal (Pa). ERP comparisons used repeated measures analyses of variance (ANOVAs) (criterion $p = .05$). Mauchley's test for sphericity was used to test for sphericity violations when analyses involve three or more levels of a repeated measure, and the Greenhouse-Geisser correction

was applied whenever nonsphericity was present. An additional 10 Hz low-pass filter was applied to grand averages for display purposes only.

Results: Behavior

Study-phase ratings—Table 1A shows the average proportion of abstract-concrete and concrete-concrete trials assigned each rating on the 6-point ease-of-definition scale during the conceptual combination encoding blocks. A comparison of the average rating assigned to concrete-concrete trials (mean = 4.00, $SE = 0.09$) versus abstract-concrete trials (mean = 3.80 $SE = 0.12$) was nonsignificant, $t(23) = 1.81$, $p = .08$, Cohen's $d = 0.37$.

Table 1B shows the average proportion of abstract-concrete and concrete-concrete trials assigned each rating on the 6-point frequency-comparison scale during the frequency-comparison encoding blocks. Average ratings did not differ between concrete-concrete (mean = 3.48, $SE = 0.06$) and abstract-concrete pairs (mean = 3.62, $SE = 0.08$, $t(23) = 1.45$, $p = .16$, Cohen's $d = 0.30$).

Associative recognition performance—To assess associative recognition memory, discrimination sensitivity (d') was calculated separately for each combination of task and W1 concreteness level. The results are depicted in Figure 2. To formally assess the effects of the task and concreteness manipulations on associative memory, a 2 (Task: compare/combine) $\times$ 2 (W1 Type: abstract/concrete) repeated measures ANOVA was performed on d' values. A significant main effect of task emerged, $F(1,23) = 105.22$, $p < .001$, partial eta-squared (η_p^2) = .82, indicating higher recognition accuracy in conceptual combination blocks relative to frequency-comparison blocks. In addition, a main effect of W1 type, $F(1,23) = 26.59$, $p < .001$, $\eta_p^2 = .54$, revealed greater accuracy for concrete-concrete relative to abstract-concrete word pairs. The Task $\times$ W1 Type interaction was not significant, $F(1,23) = 0.67$, $p = .42$, $\eta_p^2 = .03$.

Study-phase ratings and associative memory—An additional analysis assessed the relationship of participants' ease-of-definition ratings during the conceptual combination task to subsequent recognition accuracy. Data from four participants who did not register any "miss" trials on conceptual combination blocks were excluded from analysis. A paired t test indicated that the average rating given at study was significantly greater for items that were later remembered (mean rating = 4.06, $SE = 0.12$) relative to those that were later forgotten (mean rating = 2.88, $SE = .27$, $t(19) = 4.75$, $p < .001$, Cohen's $d = 1.06$). The average rating assigned during the frequency-comparison task did not significantly differ between later-remembered (mean = 3.69, $SE = 0.10$) and later-forgotten trials (mean = 3.89, $SE = 0.13$, $t(22) = 1.70$, $p = .10$, Cohen's $d = 0.35$; one participant excluded for registering no miss trials).

Results: Electrophysiology

Effects of the task and W1 concreteness manipulations on ERPs were analyzed separately for first words (W1s) and second words (W2s). The time window of 300–500 ms was chosen for N400 analyses based on extensive prior research (Kutas & Federmeier, 2011). The time window of 700–1,000 ms was chosen for N700 analyses based on a combination of visual

inspection and prior research, in which N700 latencies have varied somewhat but tend to occur in the proximity of 700 ms (Gullick et al., 2013; Holcomb et al., 1999; Huang et al., 2010; Welcome et al., 2011; West & Holcomb, 2000). Each analysis took the form of 2 (Task: compare/combine) $\times$ 2 (Concreteness: abstract/concrete) $\times$ 3 (Electrode Cluster: anterior frontal/frontocentral/parietal) ANOVAs. Effects involving electrode cluster are reported only in the context of interactions with other variables.

Task and concreteness effects on W1s—We first compared N400 and N700 ERPs to concrete and abstract W1s in the frequency-comparison and conceptual combination tasks. The resulting waveforms are depicted in Figure 3, and relevant topographical plots can be found in Figure 4. The mean number of artifact-free trials per participant per condition was 34 (range = 26–36),

300–500 ms: From 300–500 ms, a significant main effect of concreteness emerged, $F(1,23) = 11.16$, $p = .003$, $\eta_p^2 = .33$, indicating more negative amplitudes for concrete than for abstract words. No other main effects or interactions were significant, $F_s < 0.48$, $p_s > .55$. These results suggest that concreteness effects for W1s were similar across the two tasks and broadly distributed over the head.

700–1,000 ms: Despite the appearance of more negative ERPs for concrete relative to abstract W1s in the frequency-comparison task, statistical analyses revealed no significant main effects or interactions for any variables in this time window, $F_s < 1.85$, $p_s > .18$.

In summary, word-level concreteness effects in both tasks took the form of a broadly distributed increased negativity in the N400 time window.

Task and concreteness-modification effects on W2s—We next examined concreteness-modification effects present in both tasks by comparing ERPs to W2s that were preceded by concrete W1s with those that were preceded by abstract W1s. The resulting waveforms are depicted in Figure 5, and relevant topographical plots can be found in Figure 6. The mean number of artifact-free trials per participant was 32 for each condition (range = 22–36).

300–500 ms: A significant main effect of W1 concreteness emerged from 300–500 ms, $F(1,23) = 5.74$, $p = .03$, $\eta_p^2 = .20$, indicating more negative amplitudes for abstractly modified than for concretely modified words. No other main effects or interactions were significant, $F_s < 1.89$, $p_s > .18$.

700–1,000 ms: Analyses of the 700–1,000 ms window revealed a significant three-way interaction between task, W1 concreteness, and electrode cluster, $F(1.22,28.06) = 5.53$, $p = .02$, $\eta_p^2 = .22$. No other main effects or interactions were significant, $F_s < 3.41$, $p_s > .07$. A series of 2 (Task: compare/combine) $\times$ 2 (W1 Concreteness: abstract/concrete) ANOVAs were conducted for each of the three electrode clusters to further explore the three-way interaction. No main effects or interactions were significant for the parietal, $F_s < 2.35$, $p_s > .13$, or midfrontal clusters, $F_s < 3.30$, $p_s > .08$. However, analyses of the anterior frontal cluster revealed a significant main effect of W1 concreteness, $F(1,23) = 4.33$, $p = .049$, η_p^2

= .16, which was qualified by a Task $\times$ W1 Concreteness interaction, $F(1,23) = 4.42$, $p = .047$, $\eta_p^2 = .16$. Follow-up paired t tests revealed that ERPs in this cluster were significantly more negative for concretely modified versus abstractly modified W2s in the conceptual combination task, $t(23) = 3.55$, $p = .002$, Cohen's $d = 0.73$,⁴ whereas no significant difference was present in the frequency-comparison task, $t(23) = 0.26$, $p = .80$, Cohen's $d = .05$.

Study-phase ratings and ERPs—A final analysis involved subdividing ERPs from the conceptual combination task into trials that were rated as relatively difficult to define (e.g., given a rating of 1, 2, or 3) and trials that were rated as relatively easy to define (rating of 4, 5, or 6). Trials were classified separately for each participant according to his or her own ratings. We conducted 2 (Rating: easy/ hard) $\times$ 2 (Electrode Cluster: anterior frontal/ frontocentral/parietal) ANOVAs separately for W1s and W2s over the time windows of the N400 (300–500 ms) and the frontal negativity (700–1,000 ms). The only finding that emerged was a marginal main effect of rating over the N400 window for W1s, reflecting a trend for more negative N400 amplitudes for word pairs rated as easy versus hard to define, $F(1,23) = 3.32$, $p = 0.08$, all other p s $> .15$.

Discussion

The primary goal of Experiment 1 was to examine the neural mechanisms of flexible conceptual combination—in particular, the integration of arbitrary word pairs with little-to-no combinatory history. We thus analyzed ERPs to sequentially presented, unrelated noun pairs under task instructions that encouraged these pairs to be processed either as individual concepts (in a frequency-comparison task) or as modifier-head dyads (in a conceptual combination task). By systematically manipulating the concreteness of the first word of each pair (which occupies the modifier slot by English compounding rules), we were able to link task effects on ERPs to modification by examining compositional concreteness effects, or ways in which the concreteness of the first word served to influence processing of the second word.

In particular, we found that a late frontal negativity (N700) elicited by W2s was sensitive to the concreteness of the preceding W1s, but only during the conceptual combination task. Similar N700 compositional concreteness effects were found in Huang et al. (2010), in which the concreteness of the leading adjective of adjective-noun pairs was manipulated. Unlike in the present study, however, compositional concreteness effects in the study by Huang and colleagues were limited to pairs that formed clear integrated representations (such as “hilly farm” vs. “productive farm”). We suggest that the discrepancy between Huang and colleagues' findings and our own rests with the fact that we asked participants to attempt to identify a plausible meaning for each word pair. By contrast, Huang et al. (2010) asked participants to judge the “fit” of adjective-noun pairs, which arguably did not encourage compositional processing beyond that required to judge preexperimental coherence.

⁴One participant was identified as an outlier for this comparison because this participant's N700 compositional concreteness effect in the conceptual combination task was greater than three standard deviations above the sample mean. However, excluding this participant from analysis does not change the outcome, $t(22) = 3.65$, $p = .001$, Cohen's $d = 0.76$.

Together, these findings suggest that the N700 compositional-concreteness effect indexes an aspect of semantic integration that can be strategically deployed to support the online interpretation of novel concepts. As previously mentioned, N700 potentials have been linked in previous research to the use of sensory imagery at both the single word level (Gullick et al., 2013; Welcome et al., 2011; West & Holcomb, 2000) and the compositional level (Huang et al., 2010). The present findings are thus consistent with a role of compositional imagery in the online formation of novel conceptual combinations (see also Bower, 1970). Note, however, that this interpretation may seem at odds with the finding that concrete-concrete word pairs received only a nonsignificant trend toward higher ease-of-definition ratings than did abstract-concrete word pairs, as well as the finding that W2 N700 potentials did not differ between easy- and hard-to-combine pairs. We revisit these issues in the General Discussion as well as in a series of cross-experiment item-level analyses presented after Experiment 2 (see cross-experiment item analyses), where we present analyses of the lexical-level properties that predict perceived conceptual combination success.

We also report compositional concreteness effects for N400 potentials, which were smaller (less negative) for W2s that were preceded by concrete relative to abstract W1s. Unlike N700 compositional concreteness effects, N400 effects were present during both frequency-comparison blocks and conceptual combination blocks. This pattern suggests that N400 compositional concreteness effects may be stimulus rather than task driven, perhaps stemming from the tendency of concrete nouns to activate a wider range of semantic features. Indeed, this account also provides an explanation for why opposite-polarity N400 effects in both tasks were found at the lexical level (W1s) versus the compositional level (W2s, see also Huang et al., 2010). Said differently, if more features were activated for concrete W1s, one would expect both (1) more negative “baseline” amplitudes for concrete versus abstract W1s (e.g., Kounios & Holcomb, 1994), and (2) a higher likelihood of feature overlap with W2 and, hence, facilitation for W2’s processing that would result in smaller N400 amplitudes (e.g., Chiarello, Senehi, & Nuding, 1987; Forgács, Bardolph, Amsel, DeLong, & Kutas, 2015; Huang et al., 2010). Future research could examine whether such incidental semantic overlap can account for the finding that associative recognition memory benefited equally from W1 concreteness in the conceptual combination and frequency-comparison tasks.

As expected (e.g., Bower, 1970; McGee & Rob, 1980), associative recognition memory was (a) more accurate following the conceptual combination blocks than following frequency-comparison blocks, and (b) within conceptual combination blocks, more accurate for word pairs rated by the participant as easier versus harder to define. These behavioral patterns are consistent with, but not exclusive to, the notion that the type of compositional language processing engaged by the conceptual combination task led to more holistic and perhaps unitized representations in memory. Interestingly, one hallmark of holistic memories that has been proposed to distinguish them from other forms of associative memory is their inflexibility, or tendency to be retrieved either in their entirety or not at all (e.g., Parks & Yonelinas, 2015). By this account, explicit associative retrieval instructions should not be necessary for us to observe strong within-pair associations following conceptual combination. We thus reasoned that a clearer window into the organization of participants’ memory representations may be provided by an unstructured free recall task, in which

participants are simply told to recall as many stimuli as possible without any particular emphasis on within-pair associations. Therefore, in Experiment 2, we asked a separate group of participants to complete the same tasks that were used in Experiment 1, but added a written free recall task in between each study block and the following associative recognition test.

Importantly, we did not predict a difference in the overall number of words recalled between the conceptual combination and frequency-comparison tasks, as would be expected if memory were generically stronger in one task relative to the other. Rather, we specifically predicted stronger within-pair conditional recall probabilities—or the extent to which recall of one word within a pair was contingent on recall of the other—in the conceptual combination task.

Experiment 2

Method

Participants—Sixteen adults from the University of Illinois and surrounding areas (9 female, mean age = 21, range = 19–24 years) participated in the experiment and received either course credit or \$10. All were right-handed and native speakers of English.

Procedures—ERPs were not recorded in Experiment 2. The procedures were otherwise identical to those in Experiment 1, except that a free recall test was administered after each study block and prior to each recognition test. Following the 30-s distractor period at the end of each block, participants were given a blank piece of paper, and were asked to write down as many words as they could remember from the preceding study phase in any order. It was emphasized that it was not important for entire pairs to be remembered.

Analyses of the study-phase ratings and associative recognition performance in Experiment 2 yielded results that closely replicated those found in Experiment 1 despite the intervening free recall task. These data have been included as online supporting information.

Results

Free recall performance: Proportions of words recalled—Figure 7 depicts the mean proportions of W1s and W2s that were produced in the free recall test, subdivided by study task and W1 concreteness. A 2 (Task: compare/combine) $\times$ 2 (Word Position: W1/W2) $\times$ 2 (W1 Concreteness: abstract/concrete) repeated measures ANOVA was performed on the percentage of words recalled. No main effect of task emerged on the overall rate of word recall, $F(1,15) = 0.59$, $p = .45$, $\eta_p^2 = .04$, nor did a main effect of W1 concreteness, $F(1,15) = 2.27$, $p = .15$, $\eta_p^2 = .13$. However, a significant main effect of word position, $F(1,15) = 5.38$, $p = .04$, $\eta_p^2 = .26$, indicated that a higher proportion of W1 words was recalled relative to W2 words. In addition, a significant Task $\times$ Word Position interaction, $F(1,15) = 12.90$, $p = .003$, $\eta_p^2 = .46$, indicated that the study task manipulation differentially affected the recall of W1 versus W2 words. All interactions involving W1 concreteness were nonsignificant ($F_s < 2.94$, $p_s > .1$).

To further explore the interaction between task and word position, paired t tests compared the recall of W1 and W2 words within each task. The results indicated that W1 recall was significantly greater than was W2 recall on frequency-comparison blocks, $t(15) = 3.30$, $p = .005$, Cohen's $d = 0.82$, with no corresponding difference on conceptual combination blocks, $t(15) = 0.78$, $p = .45$, Cohen's $d = 0.19$.

In summary, overall rates of word recall were similar between the conceptual combination and frequency-comparison tasks. However, W1s showed a free recall advantage over W2s in the latter task, whereas free recall in the conceptual combination task favored neither W1s nor W2s. This pattern is consistent with the notion that word pairs subject to conceptual combination were represented in a relatively more holistic or unitized manner, and thus were more apt to be recalled in pairs rather than individually. To more directly test this prediction, we examined the effects of the task manipulation on within-pair conditional recall probability, or the extent to which a given word was more likely to be recalled insofar as its companion word was also recalled in a temporally contiguous manner.

Free recall performance: Within-pair conditional recall probability—The within-pair conditional free recall probabilities are depicted in Figure 8A, and Figure 8B illustrates the raw proportions of solo W1s, solo W2s, and complete pairs recalled in each task. Data were collapsed across levels of W1 concreteness, both for simplicity and because four participants recalled no words for at least one combination of word position, task, and concreteness level. To be counted as a recalled pair, both words must have been produced contiguously (e.g., one immediately following the other). We included pairs that were recalled in either the original or reversed order because several participants indicated that at study they would sometimes settle on a compound definition that reversed the target-modifier relationship implied by the original word order. One participant was excluded from analyses for recalling no W2s in the frequency-comparison condition.

As can be seen in Figure 8, the conceptual combination task was characterized by a dramatically higher rate of pair recall and lower rate of single-word recall when compared with the frequency-comparison task. Statistical analyses took the form of a 2 (Task: compare/combine) $\times$ 2 (Directionality: W1|W2/W2|W1) repeated measures ANOVA on the within-pair conditional free recall probabilities. The main effect of task was significant, $F(1,14) = 236.83$, $p < .000001$, $\eta_p^2 = .94$, confirming greater whole-pair recall in the conceptual combination task. The main effect of directionality was not significant, $F(1,14) = 0.09$, $p = .76$, $\eta_p^2 = .01$, nor was the Task $\times$ Directionality interaction, $F(1,14) = 3.67$, $p = .08$, $\eta_p^2 = .21$.

Study-phase ratings and free recall performance—Mirroring our analyses of the relationship between study-phase ratings and associative recognition memory, we calculated the average ratings assigned at study in the conceptual combination task for word pairs that were later recalled as pairs, word pairs for which only one word was later recalled, and word pairs for which neither word was later recalled. For the purpose of these analyses, pairs for which both words were recalled noncontiguously were grouped with single-word-recalled pairs. Two participants were excluded for either recalling no contiguous pairs ($n = 1$) or only contiguous pairs ($n = 1$) in conceptual combination blocks.

A one-way ANOVA with the factor recall type (pair/single/ none) yielded a significant result, $F(1.25, 16.28) = 5.08$, $p = .03$, $\eta_p^2 = .58$. Follow-up t tests indicated that the mean rating given to pairs that were subsequently remembered together (mean rating = 4.96, $SE = 0.18$) was significantly higher than that given to pairs that were later partially recalled (mean rating = 3.89, $SE = .44$, $t(13) = 2.38$, $p = .03$, Cohen's $d = 0.64$), and for which neither word was recalled (mean rating = 4.14, $SE = 0.44$, $t(13) = 4.19$, $p = .001$, Cohen's $d = 1.12$). The difference between ratings given to partially recalled pairs did not differ from (and in fact was numerically lower than) those given to nonrecalled pairs, $t(13) = 0.71$, $p = .49$, Cohen's $d = 0.19$.

Discussion

Experiment 2 provided additional insight into the memory benefits provided by the conceptual combination task used in Experiment 1. First, these benefits did not extend to all forms of memory, as evidenced by the fact that participants' ability to recall individual words was task invariant. Second, a differential reliance on holistic memory representations during conceptual combination versus frequency-comparison blocks was evident during the free recall task, despite the fact that this task did not explicitly assess memory for associations. Participants were told that the purpose of the free recall task was simply to assess how many individual words they were able to recall from each study block. Nonetheless, 82% of words recalled on conceptual combination blocks were recalled as part of a contiguous pair, compared to only 21% of words recalled on frequency-comparison blocks. This pattern suggests that the present task—much like other strategies associated with unitization (e.g., Parks & Yonelinas, 2015)—resulted in memory representations that tended to be relatively inflexible, such that they were difficult to access in a piecemeal manner.

Of course, the present behavioral patterns cannot speak to the extent to which associative memory representations formed during conceptual combination blocks ultimately came to resemble those of single items on a neural level (e.g., amenable to perirhinal-mediated retrieval), as has been found in studies of experimenter-guided conceptual combination (Haskins et al., 2008). However, the fact that item memory did not differ between the conceptual combination and frequency-comparison tasks is suggestive of a relative difference in the level of unitization afforded to the resulting memories, as opposed to a more general explanation such as levels of processing (which tends to improve both item and associative memory; Parks & Yonelinas, 2015). Accordingly, the ERP correlates of flexible conceptual combination identified in Experiment 1 can be taken as potential candidates for component processes of the self-initiated unitization of meaningful materials (although future research will be needed to support this relationship more directly, as will be discussed in the General Discussion).

One unpredicted finding was that free recall rates during frequency-comparison blocks were higher for words that were presented first in each pair (W1s) relative to those presented second in each pair (W2s). Although this result was unexpected, it is perhaps not surprising given that W1s were presented earlier in the course of each trial, and thus were afforded more processing time than were W2s. The fact that the conceptual combination task failed to

show this W1 benefit, taken together with the dramatic increase in withinpair conditional recall in the conceptual combination task, suggests that the memorability of participants' combined concepts in this task overshadowed any mnemonic benefits afforded by the temporal precedence of the W1s. This finding is consistent with our analyses of task-related changes to ERPs in Experiment 1, in which significant task-related differences were found only on ERPs to W2s.

A final question of interest concerned participants' trial-by-trial ratings of subjective ease of conceptual combination. In both experiments, word pairs given higher ease-of-definition ratings were correctly recognized and holistically recalled at higher rates than were word pairs rated as more difficult to define. We took advantage of this introspective accuracy to conduct post hoc analyses of (a) whether certain stimulus properties of W1s and/or W2s were associated with greater amenability to conceptual combination, and (b) whether combination success mediated a relationship between these stimulus properties and successful subsequent memory.

Given our proposal that conceptual combination is bolstered by compositional imagery, we were particularly interested in whether normative ratings of word imageability would be associated with higher ease-of-definition ratings.⁵ Additionally, we identified word familiarity and set size (number of near neighbors) as potential correlates of the number and availability of combinatorial possibilities associated with a given word, which may also facilitate conceptual combination (e.g., Gagné & Spalding, 2013). Using multiple regression analyses, we examined the extent to which these properties account for variance in participants' ratings when associated with words that were presented in the temporal order that is typical of the modifier (W1) or head noun (W2) of a given word pair. For W1 analyses, we also included as predictors interaction terms of each variable with W1 concreteness, which was coded as a categorical variable (0 = abstract, 1 = concrete) to reflect the bimodal distribution of concreteness within the W1 stimulus set. Stimulus properties that were significant within the regression models then served as independent variables in mediation analyses in which mean ease-of-definition rating served as the mediator and mean memorability (e.g., the mean "oldness" rating assigned when presented as an "intact" pair at test, recoded so that a rating of 6 corresponded to "confident old" and a rating of 1 corresponded to "confident new" in all participants) served as the dependent variable. Significant indirect effects in these circumstances would suggest that the stimulus property in question enhanced pair memorability specifically by way of facilitating conceptual combination.

Cross-experiment item analyses—Imageability and familiarity norms were taken from the MRC Psycholinguistic Database (Coltheart, 2007), and set size norms from the South Florida Free Association Norms (Nelson et al., 2004). Set size information was available for 51/72 abstract W1s, 58/72 concrete W1s, and 129/144 W2s, and item analyses were therefore restricted to this subset of the stimuli. The dependent measure used was the

⁵Imageability as a construct is closely related to concreteness, and the two were highly correlated in our stimulus set ($r = .88$). However, research suggests that the two concepts are not interchangeable (e.g., Dellantonio, Mulatti, Pastore, & Job, 2014), and the distribution of imageability norms among our set of modifiers was relatively continuous despite the bimodal nature of the concreteness ratings.

average rating assigned across both experiments to any word pair involving the W1 or W2 in question, regardless of its companion word. Due to the counterbalancing structure of both experiments, only half of the participants encountered each word during the conceptual combination task. Accordingly, a total of 20 participants (12 from Experiment 1 and 8 from Experiment 2) contributed to the average ratings for each word.

As shown in Table 2A, the model including W1 imageability, familiarity, set size, and the concreteness interaction terms accounted for significant variance in participants' average rating, $F(6,102) = 3.49$, $p = .004$, $R^2 = .17$. However, only W1 imageability emerged as a significant independent predictor.⁶ The Imageability $\times$ W1 Concreteness interaction was not significant, suggesting that W1 imageability was positively associated with participants' ratings for both abstract-concrete and concrete-concrete word pairs. Indeed, exploratory correlational analyses revealed that the association between average rating and imageability for abstract-concrete word pairs, $r(70) = .55$, $p < .001$, was numerically stronger than the corresponding association for concrete-concrete word pairs, $r(70) = .38$, $p = .001$.

Table 2B depicts the outcome of the regression model involving W2 attributes. As shown, the variance explained by the model as a whole was statistically significant, $F(3,125) = 3.94$, $p = .01$, $R^2 = .09$. However, the only significant independent predictor was W2 set size. The outcome of the regression analysis did not change when the dependent variable included only those ratings that were obtained when a W2 had been paired with an abstract W1, nor only ratings obtained for W2s encountered in the context of a concrete W1.

Item analyses thus provided converging evidence for a facilitative role of imagery in the conceptual combination process by demonstrating that word pairs involving more imageable W1s were perceived as easier to define relative to those with less imageable W1s. In addition, W2 semantic set size—or number of close semantic neighbors—emerged as an additional contributor to perceived amenability to conceptual combination. Follow-up mediation analyses were thus conducted for each of these stimulus properties. Mediation analyses were performed using the lavaan package in R (Rosseel, 2012). Path diagrams for both analyses are depicted in Figure 9. As shown in Figure 9A, the direct relationship between W1 imageability and pair memorability was nonsignificant, $\beta = .02$, $SE = 0.10$, $p = 0.87$.⁷ As expected, however, W1 imageability was a significant positive predictor of ease-of-definition, $\beta = .36$, $SE = 0.09$, $p < .001$, and ease-of-definition was a significant predictor of pair memorability, $\beta = .27$, $SE = 0.10$, $p = .007$. Most notably, the indirect effect in the model was also significant, $\beta = .10$, $SE = 0.04$, $p = .03$, indicating mediation. Figure 9B depicts the results of the mediation analysis with respect to W2 set size. Mirroring the findings for W1 imageability, the direct relationship between W2 set size and pair memorability was nonsignificant, $\beta = -.05$, $SE = 0.09$, $p = .56$, whereas significant path coefficients related W2 set size to ease-of-definition, $\beta = .26$, $SE = 0.08$, $p = .001$, and ease-

⁶It is, of course, important to bear in mind that null findings with respect to the other variables and interaction terms may be a consequence of a lack of statistical power, particularly given the number of variables included in the regression models.

⁷Note that according to some traditional approaches to mediation analyses—most notably the influential approach of Baron & Kenny (1986)—this lack of a significant zero-order relationship between the independent and dependent variable would preclude testing for the significance of a mediator. However, contemporary thinking about mediation (e.g., Hayes, 2009; Hayes & Rockwood, 2016; Zhao, Lynch, & Chen, 2010) imposes no such requirement.

of-definition to pair memorability, $\beta = .41$, $SE = 0.10$, $p < .001$. The indirect effect was again significant, $\beta = .11$, $SE = 0.04$, $p = .008$.

In summary, these cross-experiment item analyses provide support for a causal model by which (a) W1 imageability and W2 set size each facilitated conceptual combination success, and (b) these increases in combination success, in turn, promoted pair memorability. Below, we discuss implications of these findings, together with those of Experiment 1 and 2, with respect to the neurocognitive mechanisms of flexible conceptual combination and the circumstances under which it is most likely to serve as an effective memorization strategy.

General Discussion

An extensive body of work has demonstrated that the neural mechanisms of associative memory retrieval are not unitary. Rather, they depend on multiple factors, including individual variability in pre- and post-stimulus processing (e.g., Addante, de Chastelaine, & Rugg, 2015; Cohen et al., 2015) as well as the nature and complexity of the relevant memory representations. With regard to the latter, prior work has shown that associations that are subject to relatively higher levels of unitization at encoding will subsequently show less dependence on the hippocampal relational memory system (Haskins et al., 2008; Quamme et al., 2007). However, the field has focused less on the question of how people can strategically direct initial information processing so as to increase the amount of unitization afforded to the resulting representation. The present study addressed this gap by investigating the flexible deployment of combinatory language mechanisms to word pairs that would otherwise likely be perceived as arbitrary.

In two experiments, healthy young adult participants showed an associative recognition memory advantage after completing a flexible conceptual combination task compared to a task that did not encourage combinatory processing. Moreover, free recall performance in Experiment 2 suggested that, following the combination task, word pairs were represented in a comparably holistic and inflexible manner consistent with a higher level of unitization. Most interestingly, analyses of ERPs shed light on how flexible conceptual combination was accomplished. Specifically, we found that N700 potentials to W2s were augmented when those W2s were preceded by concrete versus abstract W1s, but only when participants were attempting conceptual combination. The present data thus extend our understanding of compositional concreteness effects—at least as reflected in N700 potentials—by illustrating that they are not limited to word pairs that correspond to preexperimentally familiar concepts. Rather, these ERPs index an aspect of conceptual combination that can be flexibly applied to a range of situations in which words are processed as modifier-head dyads, even when doing so necessitates novel or even implausible interpretations.

As previously mentioned, N700 compositional concreteness effects in this context may reflect a reliance on imagery when identifying and/or evaluating novel definitions (Gullick et al., 2013; Huang et al., 2010; Welcome et al., 2011; West & Holcomb, 2000). Indeed, item analyses revealed (a) a positive correlation between the normed imageability of a given W1 and the average ease-of-definition rating assigned to its associated word pairs, and (b) a mediation effect, whereby ease-of-definition mediated an indirect relationship between

imageability and associative memory. Together, these results lend further credence to the notion that participants engaged in imagery or visualization-based strategies when completing the conceptual combination task, and provide preliminary evidence that the N700 compositional concreteness effects evoked during this task were not epiphenomenal to the task demands. These findings are also consistent with long-standing cognitive theories of concreteness effects such as dual-coding theory (e.g., Paivio, 1969, 1991), which credits the imageability of concrete words for the many advantages they enjoy in processing and memory. Future work could test other implications of dual-coding theory, such as the notion that presenting stimuli to the left visual field/right hemisphere might increase the contributions of imagery to conceptual combination and memory success (Paivio & Allan, 2007).

Importantly, additional research will also be necessary to determine the extent and precise nature of the relationship between the N700 compositional concreteness effects that emerged during the conceptual combination task and the beneficial effects of this task on associative recognition and recall. In particular, future work could determine whether trial-by-trial variability in N700 amplitudes at study is predictive of subsequent memory at the individual trial level (e.g., “subsequent memory analyses,” Paller & Wagner, 2002). Such analyses were not possible in the present study owing to near-ceiling levels of associative recognition in the conceptual combination condition of Experiment 1, which resulted in an average of only two “subsequent miss” trials per participant.

Future work focusing on such trial-level relationships—particularly in the context of stimuli with varying levels of stimulus imageability and task emphasis on imagery—would not only help to establish more direct relationships among N700 potentials, imagery, and memory formation, but also probe the nature of this relationship. One possibility, for example, is that N700 compositional concreteness effects are diagnostic of a processing orientation toward attempting compositional imagery and/or compositional processing in general, but are relatively insensitive to the outcomes of these attempts on any given trial. Indeed, similar phenomena can be found in ERP studies of memory retrieval (e.g., Rugg & Wilding, 2000), in which certain brain potentials signal a mental set toward retrieving a particular type of information (a retrieval orientation) but are invariant with respect to retrieval success. Such a processing orientation account of the present findings are consistent with the finding that ERP comparisons of the W2s of word pairs rated as easy versus hard to define yielded no significant differences. On the other hand, the fact that N700 ERPs showed a Task $\times$ W1 Concreteness interaction rather than a main effect of task in Experiment 1 is suggestive of some form of systematic within-task variability, either in the amount of effort or cognitive resources put toward combining a given word pair, the success or outcome of the combination attempt, or some mixture of both. The present results should thus motivate additional research that will deepen our understanding of the neural processes that allow novel compounds to be interpreted and represented holistically, including the role of imagery in facilitating these processes.

Imageability was not the only stimulus property that was suggested by our cross-experiment item analyses to facilitate conceptual combination. Set size (number of near neighbors) of the W2s also emerged as both a predictor of ease of definition and, by virtue of this

relationship, an indirect predictor of pair memorability. The selectivity of this finding to W2s should be interpreted with caution given that the sets of words that served as W1s versus W2s were not counterbalanced. Nonetheless, this pattern aligns with current cognitive theories of conceptual combination, which suggest somewhat different roles for the modifier and head noun in the interpretive process. In particular, work by Gagné and colleagues (e.g., Gagné & Shoben, 1997; Gagné & Spalding, 2013) suggests that the initial stage of compound interpretation is governed largely by the perceiver's knowledge about the combinatorial history of the modifier. This knowledge is used to suggest a set of candidate relational structures, which are subsequently screened for compatibility with the head noun. For example, while the modifier "caffeine" might initially activate both the *contained in* and *causes* relations, different candidates will be selected for the head nouns "pill" versus "buzz." This account suggests that W2s with higher set sizes (e.g., "hand" or "shield") were associated with higher ease-of-definition ratings than were W2s with lower set sizes ("toad") because the former were more likely to be compatible with at least one relation proposed by the W1. Future research that specifically manipulates the number and strength of within-pair compatible relations could reveal how the neural mechanisms of conceptual combination are influenced by the difficulty of the relational matching process, both independently of and in concert with variation in concreteness or imageability.

As previously discussed, prior work has demonstrated that experimenter-guided conceptual combination can disproportionately benefit associative memory accuracy in populations who tend to experience associative memory deficits, such as older adults. The present study was motivated in part by a need to understand the processes that are involved in flexible, self-initiated conceptual combination in order to make informed decisions about when and for whom it is best suited as a memorization strategy. Of course, future work that directly examines older adults and/or other relevant populations will be required to speak conclusively to these issues. Nonetheless, at least two aspects of the present findings provide additional impetus for such work. The first comes from our item-level analyses, which link imagery to the beneficial effects of the conceptual combination task on associative memory accuracy. While much remains to be understood about the effects of aging on imagery, there is evidence that older adults are less likely to spontaneously engage in imagery during encoding tasks (e.g., Dirks & Craik, 1992; Hertzog, Price, & Dunlosky, 2012), show altered neural activity during explicit imagery instructions (Kalkstein, Checksfield, Bollinger, & Gazzaley, 2011), and may be particularly burdened by increases in the complexity of an imagery task (Dror & Kosslyn, 1994). Secondly, our finding that N700 compositional concreteness effects emerged during the conceptual combination task is also of potential relevance to the aging literature given that such effects have previously been shown to be age sensitive. Indeed, a study by Huang, Meyer, and Federmeier (2012) found that compositional concreteness effects—including N700 effects—were entirely absent in older adults during the same adjective-noun processing task that produced such effects in younger adults (Huang et al., 2010).

Together, these results underscore the need for future research to examine when and to what extent individuals across the adult lifespan can benefit from the use of self-directed conceptual combination as a mnemonic strategy. In addition to further electrophysiological research, fMRI studies could determine whether more difficult combination problems are

associated with greater recruitment of age- and disease-sensitive brain regions such as the lateral prefrontal cortex, which has been previously implicated in semantic association generation (e.g., Addis, Giovanello, Vu, & Schacter, 2014), or the hippocampus, which has been linked to performance on divergent thinking tasks not dissimilar to flexible conceptual combination (Duff, Kurczek, Rubin, Cohen, & Tranel, 2013). Such findings would speak to the question of how and when the outcomes of conceptual combination attempts may differ across the adult lifespan or in populations associated with prefrontal cortex or medial temporal lobe dysfunction.

Finally, a clear additional goal for future research is to confirm that the memory benefits afforded by the present flexible conceptual combination task did, in fact, result from increased unitization of the resulting representations at a neural level. For example, fMRI could be used in future work to examine the extent of perirhinal involvement during associative recognition and/or free recall following this task. Regardless, the present study is among the first to investigate the flexible implementation of a processing strategy that has been associated with unitization in numerous previous studies. In this way, our study provides a starting point to investigate the neurocognitive resources that allow conceptual combination to be applied in a flexible and self-directed manner in service of both language and memory tasks.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by National Institute on Aging Grant R01-AG026308, a James S. McDonnell Foundation Scholar Award to KDF, and a Beckman Postdoctoral Fellowship to HDL. We thank Kathy Abusager and Resh Gupta for assistance with data collection.

References

- Addante RJ, de Chastelaine M, Rugg MD. Pre-stimulus neural activity predicts successful encoding of inter-item associations. *NeuroImage*. 2015; 105:21–31. DOI: 10.1016/j.neuroimage.2014.10.046 [PubMed: 25450109]
- Addis DR, Giovanello KS, Vu M-A, Schacter DL. Age-related changes in prefrontal and hippocampal contributions to relational encoding. *NeuroImage*. 2014; 84:19–26. DOI: 10.1016/j.neuroimage.2013.08.033 [PubMed: 23994453]
- Ahmad FN, Fernandes M, Hockley WE. Improving associative memory in older adults with unitization. *Aging, Neuropsychology, and Cognition*. 2014; 22(4):452–472. DOI: 10.1080/13825585.2014.980216
- Bader R, Mecklinger A, Hoppstädter M, Meyer P. Recognition memory for one-trial-unitized word pairs: Evidence from event-related potentials. *NeuroImage*. 2010; 50(2):772–81. DOI: 10.1016/j.neuroimage.2009.12.100 [PubMed: 20045471]
- Baron RM, Kenny DA. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*. 1986; 51(6):1173–1182. DOI: 10.1037/0022-3514.51.6.1173 [PubMed: 3806354]
- Bastin C, Diana RA, Simon J, Collette F, Yonelinas AP, Salmon E. Associative memory in aging: The effect of unitization on source memory. *Psychology and Aging*. 2013; 28(1):275–283. DOI: 10.1037/a0031566 [PubMed: 23527745]

- Bemis DK, Pylkkänen L. Basic linguistic composition recruits the left anterior temporal lobe and left angular gyrus during both listening and reading. *Cerebral Cortex*. 2013; 23(8):1859–1873. DOI: 10.1093/cercor/bhs170 [PubMed: 22735156]
- Bower GH. Imagery as a relational organizer in associative learning. *Journal of Verbal Learning and Verbal Behavior*. 1970; 9(5):529–533. DOI: 10.1016/s0022-5371(70)80096-2
- Castel AD, Craik FIM. The effects of aging and divided attention on memory for item and associative information. *Psychology and Aging*. 2003; 18(4):873–885. DOI: 10.1037/0882-7974.18.4.873 [PubMed: 14692872]
- Chiarello C, Senehi J, Nuding S. Semantic priming with abstract and concrete words: Differential asymmetry may be postlexical. *Brain and Language*. 1987; 31(1):43–60. DOI: 10.1016/0093-934x(87)90060-5 [PubMed: 3580839]
- Cohen N, Pell L, Edelson MG, Ben-Yakov A, Pine A, Dudai Y. Peri-encoding predictors of memory encoding and consolidation. *Neuroscience and Biobehavioral Reviews*. 2015; 50:128–142. DOI: 10.1016/j.neubiorev.2014.11.002 [PubMed: 25446944]
- Coltheart M. The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology Section A*. 2007; 33(4):497–505. DOI: 10.1080/14640748108400805
- Davachi L. Item, context and relational episodic encoding in humans. *Current Opinion in Neurobiology*. 2006; 16(6):693–700. DOI: 10.1016/j.conb.2006.10.012 [PubMed: 17097284]
- Del Prato P, Pylkkänen L. MEG evidence for conceptual combination but not numeral quantification in the left anterior temporal lobe during language production. *Frontiers in Psychology*. 2014; 5:524.doi: 10.3389/fpsyg.2014.00524 [PubMed: 24926273]
- Dellantonio S, Mulatti C, Pastore L, Job R. Measuring inconsistencies can lead you forward: Imageability and the x-ception theory. *Frontiers in Psychology*. 2014; 5:708.doi: 10.3389/fpsyg.2014.00708 [PubMed: 25076920]
- Dirkx E, Craik FI. Age-related differences in memory as a function of imagery processing. *Psychology and Aging*. 1992; 7(3):352–358. DOI: 10.1037/0882-7974.7.3.352 [PubMed: 1388855]
- Dror IE, Kosslyn SM. Mental imagery and aging. *Psychology and Aging*. 1994; 9(1):90–102. DOI: 10.1037/0882-7974.9.1.90 [PubMed: 8185873]
- Duff MC, Kurczek J, Rubin R, Cohen NJ, Tranel D. Hippocampal amnesia disrupts creative thinking. *Hippocampus*. 2013; 23(12):1143–1149. DOI: 10.1002/hipo.22208 [PubMed: 24123555]
- Estes Z, Jones LL. Integrative priming occurs rapidly and uncontrollably during lexical processing. *Journal of Experimental Psychology: General*. 2009; 138(1):112–130. DOI: 10.1037/a0014677 [PubMed: 19203172]
- Forgács B, Bardolph MD, Amsel BD, DeLong KA, Kutas M. Metaphors are physical and abstract: ERPs to metaphorically modified nouns resemble ERPs to abstract language. *Frontiers in Human Neuroscience*. 2015; 9:28.doi: 10.3389/fnhum.2015.00028 [PubMed: 25713520]
- Gagné CL, Shoben EJ. Influence of thematic relations on the comprehension of modifier–noun combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1997; 23(1):71–87. DOI: 10.1037/0278-7393.23.1.71
- Gagné CL, Spalding TL. Conceptual composition: The role of relational competition in the comprehension of modifier–noun phrases and noun–noun compounds. *Psychology of Learning and Motivation*. 2013; 59:97–130. DOI: 10.1016/b978-0-12-407187-2.00003-4
- Gagné CL, Spalding TL, Ji H. Re-examining evidence for the use of independent relational representations during conceptual combination. *Journal of Memory and Language*. 2005; 53(3): 445–455. DOI: 10.1016/j.jml.2005.03.006
- Graves WW, Binder JR, Desai RH, Conant LL, Seidenberg MS. Neural correlates of implicit and explicit combinatorial semantic processing. *NeuroImage*. 2010; 53(2):638–646. DOI: 10.1016/j.neuroimage.2010.06.055 [PubMed: 20600969]
- Gullick MM, Mitra P, Coch D. Imagining the truth and the moon: An electrophysiological study of abstract and concrete word processing. *Psychophysiology*. 2013; 50(5):431–440. DOI: 10.1111/psyp.12033 [PubMed: 23445520]
- Hannula DE, Ranganath C. Medial temporal lobe activity predicts successful relational memory binding. *Journal of Neuroscience*. 2008; 28(1):116–124. DOI: 10.1523/jneurosci.3086-07.2008 [PubMed: 18171929]

- Haskins AL, Yonelinas AP, Quamme JR, Ranganath C. Perirhinal cortex supports encoding and familiarity-based recognition of novel associations. *Neuron*. 2008; 59(4):554–560. DOI: 10.1016/j.neuron.2008.07.035 [PubMed: 18760692]
- Hayes AF. Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*. 2009; 76(4):408–420. DOI: 10.1080/03637750903310360
- Hayes AF, Rockwood NJ. Regression-based statistical mediation and moderation analysis in clinical research: Observations, recommendations, and implementation. *Behaviour Research and Therapy*. 2016; Advance online publication. doi: 10.1016/j.brat.2016.11.001
- Hertzog C, Price J, Dunlosky J. Age differences in the effects of experimenter-instructed versus self-generated strategy use. *Experimental Aging Research*. 2012; 38(1):42–62. DOI: 10.1080/0361073x.2012.637005 [PubMed: 22224949]
- Holcomb PJ, Kounios J, Anderson JE, West WC. Dualcoding, context-availability, and concreteness effects in sentence comprehension: An electrophysiological investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1999; 25(3):721–742. DOI: 10.1037/0278-7393.25.3.721
- Huang H-W, Lee C-L, Federmeier KD. Imagine that! ERPs provide evidence for distinct hemispheric contributions to the processing of concrete and abstract concepts. *NeuroImage*. 2010; 49(1):1116–1123. DOI: 10.1016/j.neuroimage.2009.07.031 [PubMed: 19631274]
- Huang H-W, Meyer AM, Federmeier KD. A “concrete view” of aging: Event related potentials reveal age-related changes in basic integrative processes in language. *Neuropsychologia*. 2012; 50(1):26–35. DOI: 10.1016/j.neuropsychologia.2011.10.018 [PubMed: 22044648]
- Ishai A, Ungerleider LG, Haxby JV. Distributed neural systems for the generation of visual images. *Neuron*. 2000; 28(3):979–990. DOI: 10.1016/s0896-6273(00)00168-9 [PubMed: 11163281]
- Kalkstein J, Checksfield K, Bollinger J, Gazzaley A. Diminished top-down control underlies a visual imagery deficit in normal aging. *Journal of Neuroscience*. 2011; 31(44):15768–15774. DOI: 10.1523/jneurosci.3209-11.2011 [PubMed: 22049420]
- Konkel A, Warren DE, Duff MC, Tranel DN, Cohen NJ. Hippocampal amnesia impairs all manner of relational memory. *Frontiers in Human Neuroscience*. 2008; :2.doi: 10.3389/neuro.09.015 [PubMed: 18177494]
- Kounios J, Holcomb PJ. Concreteness effects in semantic processing: ERP evidence supporting dual-coding theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1994; 20(4):804–823. DOI: 10.1037/0278-7393.20.4.804
- Kounios J, Smith R, Yang W. Cognitive association formation in human memory revealed by spatiotemporal brain imaging. *Neuron*. 2001; 29(1):297–306. DOI: 10.1016/s0896-6273(01)00199-4 [PubMed: 11182100]
- Kutas M, Federmeier KD. Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*. 2011; 62:621–647. DOI: 10.1146/annurev.psych.093008.131123
- McGee R. Imagery and recognition memory: The effects of relational organization. *Memory & Cognition*. 1980; 8(5):394–399. DOI: 10.3758/bf03211135 [PubMed: 7442541]
- Mechelli A, Price CJ, Friston KJ, Ishai A. Where bottom-up meets top-down: neuronal interactions during perception and imagery. *Cerebral Cortex*. 2004; 14(11):1256–1265. DOI: 10.1093/cercor/bbh087 [PubMed: 15192010]
- Molinaro N, Carreiras M, Duñabeitia JA. Semantic combinatorial processing of non-anomalous expressions. *NeuroImage*. 2012; 59(4):3488–3501. DOI: 10.1016/j.neuroimage.2011.11.009 [PubMed: 22100767]
- Monti JM, Voss MW, Pence A, McAuley E, Kramer AF, Cohen NJ. History of mild traumatic brain injury is associated with deficits in relational memory, reduced hippocampal volume, and less neural activity later in life. *Frontiers in Aging Neuroscience*. 2013; 5:41.doi: 10.3389/fnagi.2013.00041 [PubMed: 23986698]
- Nelson DL, McEvoy CL, Schreiber TA. The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*. 2004; 36(3):402–407. DOI: 10.3758/bf03195588

- Paivio A. Mental imagery in associative learning and memory. *Psychological Review*. 1969; 76:241–263. DOI: 10.1037/h0027272
- Paivio A. Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*. 1991; 45:255–285. DOI: 10.1037/h0084295
- Paivio, Allan. *Mind and its evolution: A dual coding theoretical approach*. Mahwah, NJ: Lawrence Erlbaum Associates; 2007.
- Paller KA, Wagner AD. Observing the transformation of experience into memory. *Trends in Cognitive Sciences*. 2002; 6(2):93–102. DOI: 10.1016/s1364-6613(00)01845-3 [PubMed: 15866193]
- Parks CM, Yonelinas AP. The importance of unitization for familiarity-based learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition*. 2015; 41(3):881–903. DOI: 10.1037/xlm0000068
- Quamme JR, Yonelinas AP, Norman KA. Effect of unitization on associative recognition in amnesia. *Hippocampus*. 2007; 17(3):192–200. DOI: 10.1002/hipo.20257 [PubMed: 17203466]
- Rosseel Y. lavaan: An R package for structural equation modeling. *Journal of Statistical Software*. 2012; 48(2):1–36. DOI: 10.18637/jss.v048.i02
- Rubin RD, Chesney SA, Cohen NJ, Gonsalves BD. Using fMR-adaptation to track complex object representations in perirhinal cortex. *Cognitive Neuroscience*. 2013; 4(2):107–114. DOI: 10.1080/17588928.2013.787056 [PubMed: 23997832]
- Rugg, Wilding. Retrieval processing and episodic memory. *Trends in Cognitive Sciences*. 2000; 4(3): 108–115. DOI: 10.1016/s1364-6613(00)01445-5 [PubMed: 10689345]
- Welcome SE, Paivio A, McRae K, Joanisse MF. An elec-trophysiological study of task demands on concreteness effects: Evidence for dual coding theory. *Experimental Brain Research*. 2011; 212(3): 347–358. DOI: 10.1007/s00221-011-2734-8 [PubMed: 21656220]
- West WC, Holcomb PJ. Imaginal, semantic, and surface-level processing of concrete and abstract words: An electrophysiological investigation. *Journal of Cognitive Neuroscience*. 2000; 12(6): 1024–1037. DOI: 10.1162/08989290051137558 [PubMed: 11177422]
- Williams LE, Must A, Avery S, Woolard A, Woodward ND, Cohen NJ, Heckers S. Eye-movement behavior reveals relational memory impairment in schizophrenia. *Biological Psychiatry*. 2010; 68(7):617–624. DOI: 10.1016/j.biopsych.2010.05.035 [PubMed: 20655509]
- Zhao X, Lynch JG, Chen Q. Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*. 2010; 37(2):197–206. DOI: 10.1086/651257

A) Study Trial – Conceptual Combination **B) Study Trial – Frequency-Comparison** **C) Test Trial – Associative Recognition Test** **Figure 1.**

Sample study trial for the conceptual combination task (A) and frequency-comparison task (B), as well as a sample intact trial from the associative recognition test (C).

Figure 2.

Associative recognition memory in Experiment 1, as indicated by discrimination sensitivity (d'), subdivided by task and word pair type. Abs-con = abstract-concrete word pairs; Con-con = concrete-concrete word pairs.

Figure 3.

Grand-averaged ERPs to concrete (e.g., "skate") and abstract (e.g., "alias") first words (W1s) in Experiment 1. Waveforms are shown from 10 electrode sites, which comprise the anterior frontal and parietal clusters submitted to statistical analyses (see experimental procedures). Electrode positions are indicated on the head diagrams with solid black dots. The hollow black dots represent the frontocentral electrode cluster, which was also analyzed. A: ERPs for W1s in the frequency-comparison task. B: ERPs for W1s in the conceptual combination task.

Figure 4.

A: Grand-averaged waveforms for abstract and concrete first words (W1s) in the frequency-comparison and conceptual combination tasks at midline anterior frontal electrode Fpz and midline parietal electrode Pz. Gray shading indicates the time windows used in N400 and N700 analyses. B: Topographical plots of concreteness effects over the time windows of interest in the frequency-comparison and conceptual combination tasks.

Figure 5.

Grand-averaged ERPs to concretely modified and abstractly modified second words (W2s) in Experiment 1. Waveforms are shown from 10 representative electrode sites whose positions are indicated on the head diagrams. A: ERPs for W2s in the frequency-comparison task. B: ERPs for W2s in the conceptual combination task.

Figure 6.

A: Grand-averaged waveforms for abstractly modified and concretely modified second words (W2s) in the frequency-comparison and conceptual combination tasks at midline anterior frontal electrode Fpz and midline parietal electrode Pz. Gray shading indicates the time windows used in N400 and N700 analyses. B: Topographical plots of concreteness-modification effects over the time windows of interest in the frequency-comparison and conceptual combination tasks.

Figure 7.

Mean proportion of first words (W1s) and second words (W2s) that were produced in the free recall task during frequency-comparison blocks and conceptual combination blocks. The overall number of words recalled did not differ by task.

Figure 8.

A: Within-pair conditional recall probability in each task, or the extent to which one word from a given pair was more likely to be recalled insofar as its companion word was recalled. $W1|W2$ = the probability that a first word (W1) was recalled given that its corresponding second word (W2) was recalled; $W2|W1$ = the probability that a W2 was recalled given that its corresponding W1 was recalled. B: The raw proportion of studied pairs for which both words were recalled (teal bars), for which only the W1 was recalled (dark gray bars), or for which only the W2 was recalled (light gray bars).

Figure 9.

Path diagrams of the total and direct effects of W1 imageability (A, above) and W2 set size (B, below) on pair memorability in the combination task, as well as the indirect effects of each stimulus property with mean ease-of-definition included as a mediator.

Table 1

Distributions of Ease-of-Definition and Relative Frequency Ratings Given to Abstract-Concrete and Concrete-Concrete Trials in Experiment 1

A.	Harder to define			Easier to define		
	1	2	3	4	5	6
Abs-con	0.12 (.02)	0.16 (.02)	0.16 (.02)	0.16 (.02)	0.18 (.02)	0.22 (.03)
Con-con	0.10 (.02)	0.14 (.01)	0.11 (.01)	0.18 (.02)	0.23 (.02)	0.24 (.02)

B.	First word more frequent			Second word more frequent		
	1	2	3	4	5	6
Abs-con	0.12 (.02)	0.20 (.02)	0.16 (.02)	0.16 (.01)	0.20 (.02)	0.16 (.02)
Con-con	0.20 (.02)	0.18 (.01)	0.11 (.01)	0.15 (.01)	0.20 (.01)	0.17 (.02)

Note. Ratings are assigned on (A) the 6-point ease-of-definition scale in the conceptual combination task, or (B) the 6-point relative frequency judgment in the frequency-comparison task. Standard errors are in parentheses. abs-con = abstract-concrete; con-con = concrete-concrete.

Table 2

Results of Regression Models Investigating Properties of Words Associated with Higher Ease-of-Definition Ratings During the Conceptual Combination Task

A. First words (W1s)				
Model	<i>b</i>	<i>SE b</i>	β	<i>p</i> value
Imageability (IMG)	.002	.001	.259	.02*
Familiarity (FAM)	.001	.001	.163	.11
Set size (SET)	-.003	.012	-.029	.77
Concreteness $\times$ IMG	-.003	.002	-.120	.26
Concreteness $\times$ FAM	.001	.002	.032	.74
Concreteness $\times$ SET	.029	.023	.012	.22

B. Seconds words (W2s)				
Model	<i>b</i>	<i>SE b</i>	β	<i>p</i> value
Imageability (IMG)	.002	.002	.089	.34
Familiarity (FAM)	.001	.001	.069	.46
Set size (SET)	.033	.011	.269	.002**

Note. *b* = unstandardized coefficient; *SE b* = standard error of the unstandardized coefficient; β = standardized coefficient.