**Methods**

**Participants**

Thirty-one University of Southern Mississippi undergraduates participated for partial course credit. Three participants were excluded for failure to report 10% or more of their JOL responses (described below), leaving 28 participants for analysis. All participants were native English speakers with normal or corrected-to-normal vision. A sensitivity analysis conducted with *G\*Power* (Faul, Erdfelder, Lang, & Buchner, 2007) indicated that our sample size provided adequate power (.80) to detect a small effect size (Cohen’s *d* = 0.27) or larger.

**Materials**

One-hundred-eighty associative word pairs were taken from the University of South Florida Free Association Norms (Nelson et al., 2004). These pairs consisted of 40 asymmetric forward pairs in which association only occurred in the forward direction (e.g., bounce-ball), 40 asymmetric backward pairs in which association only occurred in the backward direction (e.g., ball-bounce), 40 symmetric pairs in which forward and backward strength were equivalent (e.g., on-oﬀ), 40 unrelated pairs (e.g., building-cat), and 20 non-tested buffers to control for primacy and recency effects. Pairs were equally distributed across two study lists, each consisting of 20 symmetrical, forward, backward, and unrelated pairs and 10 buﬀers. All participants were presented with both lists which were separated into two study-test blocks, the order of which was counterbalanced across participants. Both study lists were organized such that ﬁve buﬀer pairs were presented at the beginning and end of each list, with the remaining pairs randomized anew for each participant. Thus, each study block contained 90 pairs (80 tested, 10 buffer). Additionally, pair types were equated on associative strength (i.e., FAS and BAS) using the Nelson et al. (2004) free-association norms and lexical and semantic properties including word length, SUBTLEX frequency (Brysbaert & New, 2009), and concreteness values from the English Lexicon Project (Balota et al., 2007). Associative strength and semantic/lexical properties of the pair types are reported in the Appendix (Tables A1-A2). Furthermore, all study blocks were matched on these properties so that mean associative overlap and lexical/semantic properties were equivalent between direction types and across study lists. For all pair types, counterbalanced versions of the study lists were created that switched the order of the word pairs (i.e., forest-tree vs. tree-forest). This allowed for greater control of item differences, particularly on forward and backward pairs, as the same items were used in both the forward and backward directions across counterbalances. Pair order was similarly flipped and counterbalanced across unrelated and symmetrical pairs.

The cued-recall test in each block consisted of all 80 cues from the original study items (minus buffers). The cue was presented next to a blank space that was to be completed with the studied target word. Test order was randomized anew for each participant.

**Procedure**

All participants were tested individually via computers running *E-Prime 3* software (Psychology Software Tools, Pittsburgh, PA). Participants were instructed that they would view a series of cue-target word pairs in which the cue was always presented on the left and the target on the right and that their memory for the target word would be tested. In addition to studying the pairs, participants were instructed to provide a JOL rating. Speciﬁcally, they were told to rate the likelihood that they would be able to remember the target word in the presence of the cue word at test using a 0 to 100 scale in which 0 indicated that they would be unable to correctly recall the target word, while a response of 100 indicated full certainty that they would recall target word. Participants were encouraged to use the full range of the scale when making their judgments to limit anchoring on extremes (i.e., judgments of 0 and 100). Following instruction, participants were presented with the ﬁrst study list. The study phase was self-paced with participants viewing an item pair and typing a JOL rating before proceeding to the next pair. Participants provided JOL ratings while the pair was displayed.

Following the ﬁrst list, participants completed an arithmetic ﬁller task for 2 minutes followed by a cued-recall test in which participants were presented with the cue word from each study pair and asked to type the target word from memory. If participants were unable to retrieve the target word, they could skip to the next test cue by pressing the enter key. After completing the ﬁrst cued-recall test, participants began the second study/test block which used the same instructions as the ﬁrst block. After completion of the second study/test block, participants were fully debriefed. Each experimental session lasted approximately 30 minutes.

*Figure 1.* Comparison of mean JOL ratings and recall rates across each pair type for the item-specific encoding. Error bars represent 95% confidence intervals.

*Figure 2*. Comparison of mean JOL ratings and recall rates across each pair type for the relational encoding. Error bars represent 95% confidence intervals.

*Figure 3*. Comparison of mean JOL ratings and recall rates across each pair type for the read only control group. Error bars represent 95% confidence intervals.