A new tool for measuring individual differences in semantic categorization

Emilio Gagliardi & Chris Westbury
Department of Psychology
University of Alberta, Edmonton, AB, Canada

Introduction

Background

In order to describe and understand the structure and organization of semantic memory, investigators have often used semantic decision tasks. Here we report on a methodology that extends the traditional semantic categorization task by allowing us to measure how individual subjects associate a set of concepts.

We looked at the well-studied category of concreteness. Concrete and abstract word processing has been addressed by two influential theories. The first is Paivio's (1986) **dual code theory**, which suggests that concrete words benefit from having both a verbal and a sensory component, while abstract words have only a verbal component. Dual code theory predicts (and much evidence shows) that concrete words will be processed faster than abstract words and with fewer errors in semantic categorization experiments.

Schwanenflugel and Stowe (1989) proposed the **context availability theory**, which says that the reaction time difference between abstract and concrete words can be explained by stimulus context. They argue that "abstract words are comprehended more slowly when presented in isolation because the reader is experiencing difficulty retrieving the relevant prior knowledge" (Schwanenflugel and Stowe). They demonstrated that the performance superiority of concrete words over abstract words could be eliminated when abstract words were presented in a meaningful context.

The complexity of semantic processing might be more fully explored if the stimulus context could be measured in a way that would allow individuals to express their own associations among a set of concepts. Such a paradigm has been developed, although for very different reasons.

Initial design

The paradigm, **implicit concept mapping** (iCmap; Aidman & Egan, 1989; Aidman & Ward, 2002) is a computer-based knowledge representation methodology developed to both quantify students' understanding of a particular topic, and provide a visual re-construction of the implicit semantic structure by generating hierarchical cluster diagrams (see Figure 6).

An iCmap trial consists of three forced-choice semantic categorizations. One word is categorized as being *similar* to a target word, and two words are categorized as being *contrastive*. The process is repeated until all words have taken their turn as the target.

- Results from Aidman & Ward (2002) failed to correlate with course grade
- Results from a pilot study failed to demonstrate sensitivity to learning
- Required manual scoring of cluster diagrams by content expert
- Insufficient theoretical motivation and experimental control

Experiment

Purpose

Improve the design and methodology of the iCmap paradigm in order to extend the traditional semantic categorization task to generalize from words presented in isolation to words presented within a larger stimulus context.

Modified Design

Progressive concept mapping (proCmap) is a semantic categorization task in which participants are presented with a list of words in the center of the screen, and a target word above. The task is to judge the degree of relatedness between each word in the list and the target word by adjusting a thermometer-like rating scale below each word (see Figure 1). Each word in the list takes its turn as the target word, and when all the words have been the target, the block ends. The task is progressive because each block adds additional words to the list, starting with three words (i.e., the practice block) and ending with nine.

Each trial generates a vector of *continuous* relatedness ratings ranging from -1 (i.e., absolutely not related) to 1 (i.e., absolutely related). A complete block produces a non-symmetrical matrix of proximity data indicating the relatedness of all words to each other.

We assessed the proCmap procedure in a 2 (abstract vs. concrete) x 2 (random vs. structured) x 4 (3,5,7 & 9 words) mixed factors design with a repeated measure on the last factor. The first factor was generated by selecting words according to imageability ratings from Coltheart (1981). The second factor was created by randomly generating a list of words (controlled on several lexical dimensions) for the *random* condition, and by replacing one word in each of the random blocks so that it would have some semantic association with other words in the list, these modified lists formed the *structure* condition.

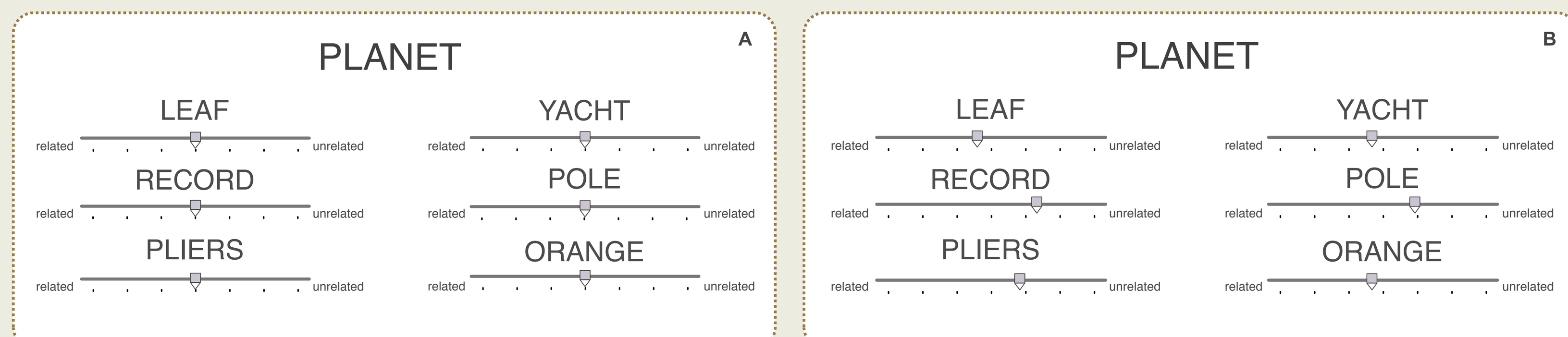

Figure 1
Graphical representation of the initial task configuration (A), which establishes the null hypothesis because each rating scale is set to zero in the middle. Panel (B) represents a completed trial from data taken from the present experiment. The relatedness between LEAF and PLANET is judged to be about 0.185 compared to about -0.35 between POLE and PLANET.

Hypotheses

This proximity data is used to produce two measures of semantic categorization:

- **Categorization complexity** (CC), which is a measure of the absolute amount of semantic information captured within a block
- **Categorization consistency** (CI), which is the average correlation of the relatedness scores within a block.

H_1: Abstract words will be rated with significantly higher CC compared to concrete words

H_2: Abstract words will be more difficult to consistently assign the same relatedness values compared to concrete words, resulting in significantly lower CI scores.

H_3: A linear effect of CC scores for the number of words in blocks two through four

H_4: A linear effect of CI scores for the number of words in blocks two through four

Results & Discussion

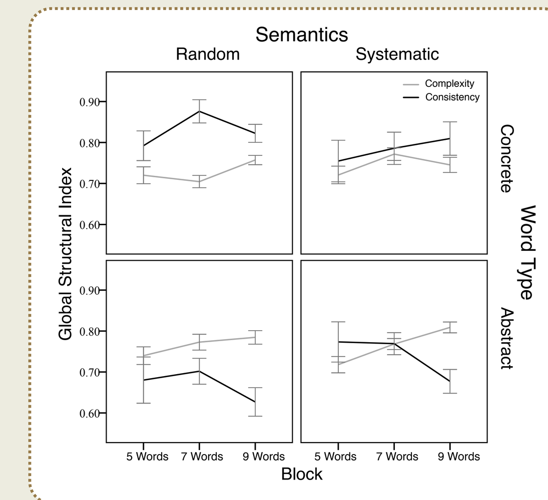

Figure 2
N=120. Performance of cognitive complexity and cognitive consistency for each condition

120 (78 females) U of A students participated for partial course credit

Hypotheses

H_1: Abstract words were judged with more complexity than concrete words (see Figure 2)
$F_{(1,115)} = 4.48$, $p = 0.036$

H_2: Abstract words are rated with less consistency than concrete words (see Figure 2)
$F_{(1,115)} = 15.78$, $p < 0.01$

H_3: There is no main linear effect for the number of words on CC
$F_{(1,115)} < 1$

H_4: There is no main linear effect for the number of words on CI
$F_{(1,115)} < 1$

Significant ANOVA effects

CI: Word type by Number of words (see Figure 3)
$F_{(1,115)} = 4.05$, $p = 0.047$

CI: Word type by Semantics (see Figure 4)
$F_{(1,115)} = 6.72$, $p = 0.011$

CC: Word type by Semantics by Number of words (see Figure 5)
$F_{(1,115)} = 3.45$, $p = 0.044$

Post-hoc analyses

CC: Semantics by Number of words quadratic interaction (see Figure 5)
$F_{(1,115)} = 7.9$, $p < 0.01$

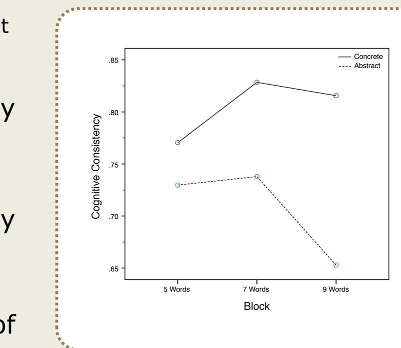

Figure 3
Word type by Number of words interaction on consistency

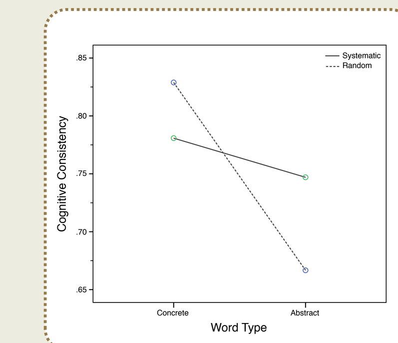

Figure 4
Word type by Semantic interaction on consistency

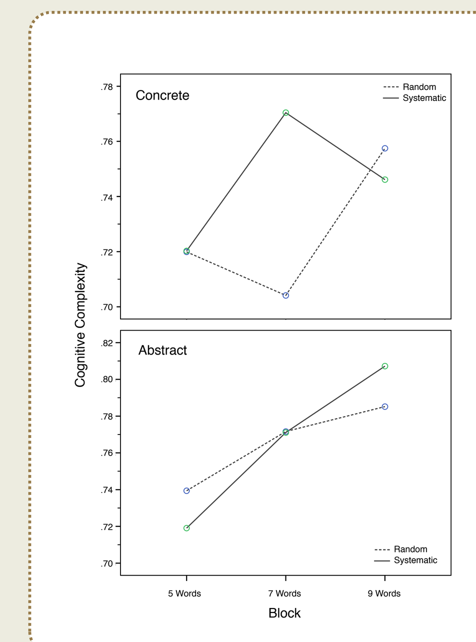

Figure 5
Word type by Semantics by Number of words interaction on complexity

Discussion

Significant effects of word type provide evidence that the proCmap procedure is sensitive to the imageability of a word's referent. The failure to demonstrate linear effects for the number of words presented simultaneously is qualified by significant 2- and 3-way interactions indicating several interesting patterns. The first is that cognitive consistency has a positive linear trend for concrete words as more words are added, and a negative linear trend for abstract words. This suggests that the salient sensory component of concrete words facilitates consistent judgements, whereas additional abstract words inhibit consistent judgments as more characteristic features are activated and compared among the concepts.

The second pattern is that cognitive consistency for abstract words is improved by adding words with meaningful associations to the stimulus context. That is, in a completely random context, concrete words outperformed abstract words, but in a meaningful context this effect was attenuated. This result provides evidence inline with Schwanenflugel and Stowe's (1989) context availability theory.

Finally, the 3-way interaction for CC indicates two further trends. First, regardless of word type, adding more words to the stimulus context increased complexity scores. Second, the quadratic trend for concrete words in a meaningful context peaked at 7 words, whereas this trend is inverted in the random condition and is greatest at 9 words. This trend was not predicted, and it did not emerge for abstract words, which suggests that in the random condition, participants initially adopted a strategy that assumed little semantic relatedness, and then switched.

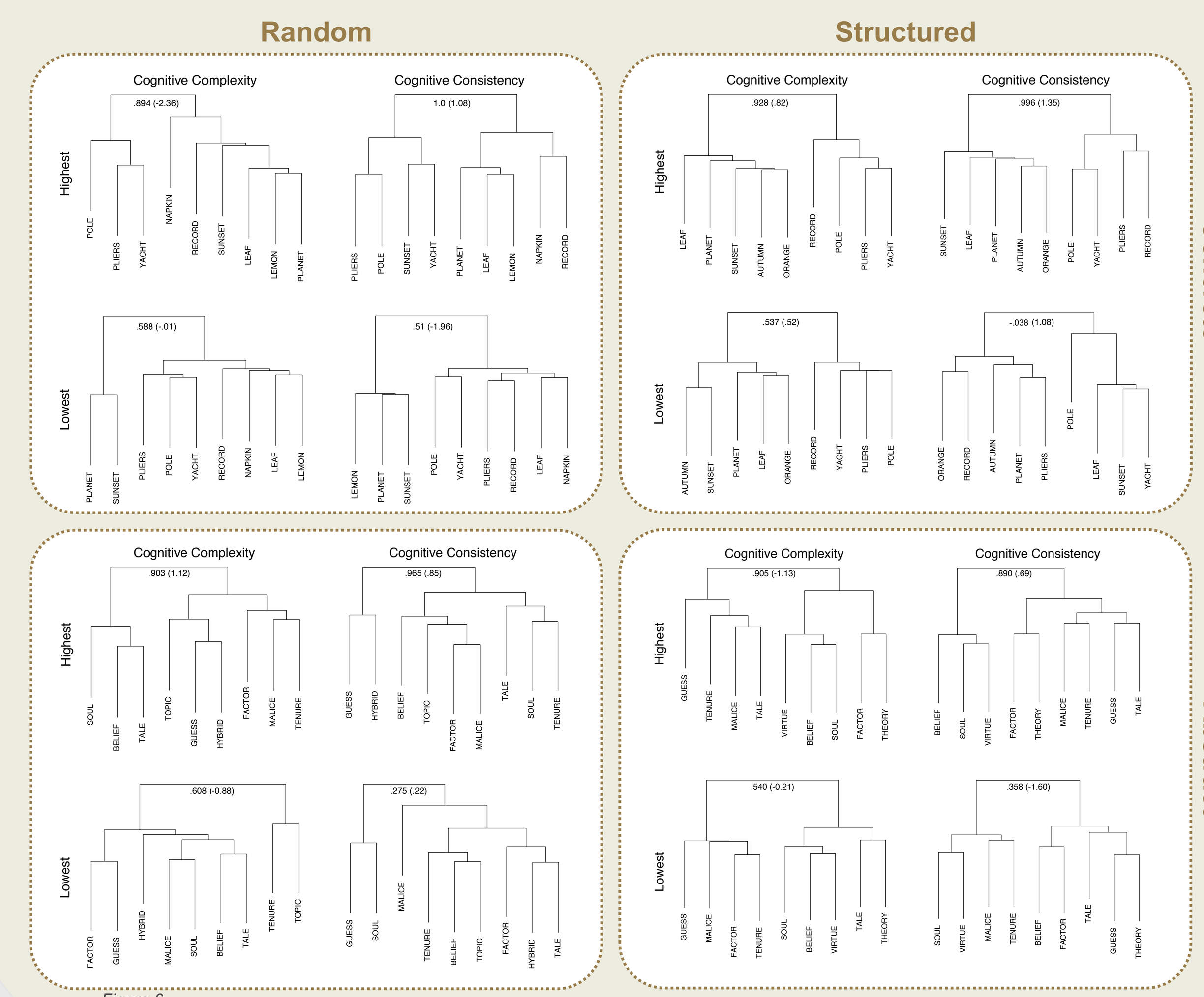

Figure 6
Cluster plot for each condition presenting the highest and lowest performing participant along with the respective score on the dependent measure. The value in parenthesis is a standardized reading comprehension score obtained prior to completing the proCmap procedure.

Cluster Plot Analysis

The cluster plots display the solutions for the participants who scored highest and lowest on both measures across the four conditions.

The vertical distance between concepts as well as cluster membership now provide formally specified information regarding the underlying semantic structure imposed by each participant. The farther away two concepts are, the less similar were they judged to be related to one another.

This is analogous to Collins and Loftus' (1975) description of the strength of the associative links between concept nodes in their spreading activation model, where shorter links represented stronger association.

Interestingly, there was no effect of reading comprehension on either measure (the value in parenthesis). However, there is a general trend where a participant scoring higher on the reading test, tended to have a higher CC or CI score.

The cluster plots were originally intended to be presented to students so that they could interpret their implicit semantic structure, and identify incorrect or superficial relationships among the concepts by comparing their 'maps' to expert maps.

In the present context, these plots help visualize the multidimensional nature of the categorization scores, and provide a data structure which captures semantic information not available in reaction time scores obtained from traditional semantic categorization experiments.

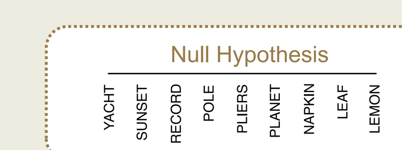

Figure 7
Cluster plot representing the null hypothesis