MEMORY ORGANIZATION OF OBJECTS BY SHAPE SIMILARITY: CLUSTERING AND INTERRESPONSE TIME EFFECTS IN FREE EMISSION AND FREE RECALL

PAUL MATTLESS

MEMORY ORGANIZATION OF OBJECTS BY SHAPE SIMILARITY: CLUSTERING AND INTERRESPONSE TIME EFFECTS IN FREE EMISSION AND FREE RECALL

by

© Paul Mattless, 2008

A thesis submitted to the School of Graduate Studies

in partial fulfilment of the requirements for the degree of

Master of Science,

Department of Psychology, Faculty of Science,

Memorial University,

St. John's, Newfoundland,

May 26, 2008.

Abstract

According to some perceptual theories of long-term memory and knowledge, object representations are organized according to shape similarity and other perceptual schemes. A prediction derived from the perceptual theories is that participants' free emission and free recall retrieval sequences should show clustering of object nouns by shape category, with shorter interresponse times for shape cluster (e.g., snake, rope) as compared to shape switch (snake, globe) transitions. However, some amodal theories state that such effects should not occur. The free emission and free recall results supported the perceptual theories, with significant shape clustering, a shape cluster speed advantage in interresponse times, and strategies and mnemonics (reported post-task) that included perceptual similarities and relations. A neural explanation, based in part on Hebb's (1949) ideas and on recent neuroscientific evidence, is proposed to account for the results.

Acknowledgements

First of all, I would like to thank my supervisor, Dr. Rita Anderson. I thank Rita for allowing me to undertake this project, for her contributions to it, and for her patience throughout. I am also fortunate that panel members Dr. Carolyn Harley and Dr. Mary Courage contributed their expertise as well. I thank Maureen Howard and Brenda Crewe, whose official administrative work also permitted this project to be carried out. I thank the members of the ICEHR (ethics committee) for reviewing the proposal for this project.

I thank all of the students who participated in the studies reported in this thesis. (Payment for the aforementioned participants was provided through an NSERC grant to Dr. Rita Anderson). Thanks also go to the students of Cognition 3450 (in 2000), who participated in pilot studies. I am also grateful to several friends and fellow psychology students, who helped with various aspects of the pilot work, including James Drover, Rochele Oliver, Roxanne Vernescu, Christa Vincent, Craig Williams, and many others who were participants in the pilot work. Thanks also go to Delf Hohmann, who translated the Zimmer (1989) methods sections from German.

Other researchers have helped me through correspondence. Dr. Larry Barsalou gave me encouragement, advice, and made constructive comments on an earlier draft of a proposal for this thesis. To my various queries, I have received helpful, detailed replies from Drs. Irving Biederman, Stephen Frey, Stephen Kosslyn, Martin Murphy, Diane Pecher, Zenon Pylyshyn, Cary Savage, and Hubert Zimmer.

In my time here at Memorial, Dr. Bill McKim has been friendly and helpful. It was a pleasure to work as a TA for him and his students. Dr. Russ Adams and Dr. John Evans have helped me get through some of the challenges that can arise, occasionally, for a graduate student within a University.

I am grateful to my mother, Barb Aiken, who has provided support morally and financially. My father, Jim Mattless, and my mother, both encouraged my boyhood interest in nature, and this probably contributed to my life-long interest in science. My brothers, Steve and John, and my sister, Debbie, have also helped me in various ways. My uncle, Fred Britton, and our friend Nick Plavac, have stimulated my thinking through our in-depth discussions exploring many areas of psychology. Finally, I thank my landlord and friend, Dr. Roger Veysoglu, for his generosity and advice.

Table of Contents

List of Tables

Note: For simplicity some of the sample size, statistical error, and other notes were removed from the above list of table titles.

List of Figures

Note: For simplicity some of the sample size, statistical error, and other notes were removed from the above list of figure captions.

List of Abbreviations and Symbols

AC	anterior cingulate cortex
ARC	Adjusted Ratio of Clustering
BG	basal ganglia
DM	dorsomedial (part of the thalamus)
DMS	delayed matching-to-sample
EEG	electroencephalography
ERP	event-related potential
FE	free emission
fMRI	functional magnetic resonance imaging
FR	free recall
IRT	interresponse time
IT	inferior temporal cortex
Ln	natural or base *e* logarithm
LOC	lateral occipital complex
lPFC	lateral prefrontal cortex
LTD	long-term depression
LTM	long-term memory
LTP	long-term potentiation
MFM	medial frontal motor cortex
mOFC	medial OFC
MRR	Modified Ratio of Clustering
MTL	medial temporal lobe

N IRTs	Number of IRTs in an output list
NR	normalized rank
OFC	orbitofrontal cortex
ROP	rank output position
PC	parietal cortex
PFC	prefrontal cortex
PNC	posterior neocortex
PP	posterior parietal cortex
pre-SMA	pre-supplementary motor area
RR	Ratio of Repetition
SO-ARC	Subjective Organization Adjusted Ratio of Clustering
STD	short-term depression
STM	short-term memory
STP	short-term potentiation
STSD	short-term synaptic depression
STSE	short-term synaptic enhancement
VIF	variance inflation factor
WM	working memory

List of Appendices

Chapter 1. Introduction

1.1. A Brief Introduction to Shape

In our everyday experience, shapes are all around us. Every physical object or entity that we know has shape. When we focus on an object or entity, much of what we see is shape. Our successful interaction with the entities and objects of the world depends on having neural systems that can capture and organize shape information effectively. Our ability to process shape information effectively allows us to accomplish a range of critical tasks: identifying foods; classifying predators and prey; distinguishing family, friends, and foes; selecting potential mates; comprehending gestures and facial expressions; recognizing everyday landmarks; reading; and so on. Knowing the basic physical properties of shapes is essential in conceiving, making, and using all kinds of objects, such as tools and utensils, shelters, clothing, vehicles, weapons, and communicative symbols. If our ability to process visual shape information is disrupted severely due to neurological damage or dysfunction, there can be severe consequences, including inability to recognize the people and things that are important to us (for reviews, see Farah, 1990; Humphreys & Forde, 2001). Shape information is not only important practically, biologically, and socially, but it is a fundamental aspect of art and design, mathematics, and the sciences.[1]

1.2. Overview and Intentions

The goal of this thesis is to explore the role of shape information in the organization of adults' knowledge or long-term memory (LTM) for everyday objects. Previous research has established that object shape plays a major role in the development of young children's object knowledge (see below). To date, very little research has explored this topic in adults.

Broadly speaking, shape appears to be implicated in the organization of our object memory according to two basic schemes. One scheme involves *relational* organization, wherein thinking of one thing brings to mind other related information. For example, hearing the word *apple* could bring to mind information such as its shape, the locations where it is found, the actions in which it is used, its taste, other objects that often occur with it, and other properties and associations. The other basic scheme occurs in *similarity* organization, wherein thinking of one thing can remind us of another that is similar. For example, you may think of a word, and then another comes to mind that sounds like it. Or you may see someone in the distance whose appearance, momentarily, reminds you of a friend. I claim that the same principle applies to object shape. If we perceive, remember, or think of an object that we know, such as an apple, we activate a shape representation that, at least momentarily, partially activates other object representations that overlap in shape and other features. Thinking of an apple could remind us of a globe, a ball, and other things having a similar shape. This tendency, when combined with relational organization, allows us to make useful generalizations about objects.[2] With experience, we can infer that a spherical object will roll; a hard object that is pointed with a sharp edge could cut tissue, and therefore should be handled carefully; a large flat object may be used as a work surface; and unfamiliar round things growing on a tree may be edible. With these combined similarity and relational organizational schemes established in the brain, seeing or knowing the shape of an object can often tell us a great deal about that object.

The primary focus of this thesis is on shape similarity organization. Some cognitive theories can be divided into two classes depending on their claims about the role of object shape similarity in memory organization. According to *perceptual* theories (Barsalou, 1999; Hebb, 1949), object memories are organized in such a way that remembering or thinking of an object, in its absence, can bring to mind other objects that are similar in shape. In extending this view, I suggest that when a person thinks of a baseball, for example, neurons tuned through experience to fire selectively for spheroid shapes become active, partially activating or increasing the potential for activation of other object representations that are implemented by many of those same neurons. In contrast, theories that can be classified as non-perceptual or *amodal* with respect to memory organization (Engelkamp & Zimmer, 1994; Shelton & Caramazza, 2001) state that object memory representations are, generally, not organized according to shape or other perceptual similarities.

In order to test which of these two types of theories is most consistent with the evidence, this research examined participants' continuous retrieval of object names in free emission and free recall tasks. In *free emission*, the task is to list as many items as one can of a certain kind within a set time limit. In *free recall*, the task is to recall as many items as possible from a previously studied set of items. Examining the retrieval output sequences from these tasks can give us insights into how the activation of one object memory (e.g., ball) influences the activation of the next one reported (e.g., apple), and so on. Both tasks are discussed in more detail in chapter 2.

This thesis is intended to make four main contributions to our understanding of human memory organization: (1). It gives the first major review focussing specifically on the perceptual similarity and relational organization of items in free emission and free recall. (2). It provides empirical evidence demonstrating which one of the two competing classes of memory theories' claims is most consistent with the evidence, thereby advancing the basic science of memory. (3). It involves the development of new experimental tasks, tasks that may, in the future, serve as useful tools in helping us learn more about memory organization in a variety of populations. (4). It lays out a neurally-based theory that attempts to explain shape similarity organization effects in free emission and free recall, extending the ideas of previous researchers in light of new evidence (chapter 3). This thesis is part of a project intended to shed light on how object and shape information, and information generally, is organized in memory.

1.3. The Role of Shape in the Development of Object Concepts

In the first few years of a child's life, shape and the shapes of parts play major roles in object classification, in the learning of object names, and in making generalizations about object properties and functions (Baldwin, 1989, 1992; Diesendruck & Bloom, 2003; Eimas & Quinn, 1994; Graham, Kibreath, & Welder, 2004; Imai, Gentner, & Uchida, 1994; Landau, Smith, & Jones, 1988; Poulin-Dubois, Frank, Graham, & Elkin, 1999; Quinn, Eimas, & Tarr, 2001; Rakison & Butterworth, 1998; Samuelson & Smith, 2000). Young children continue to be influenced by shape when they classify objects, even in some cases where researchers have tried to override shape's influence (e.g., see Gelman & Markman, 1986).

Particularly striking evidence of the role of shape in the development of object concepts can be found in young children's figurative language (Gardner, Winner, Bechhofer, & Wolf, 1978; Gelman, Croft, Fu, Clausner, & Gottfried, 1998; Kay & Anglin, 1982; Winner, McCarthy, & Gardner, 1980). A toddler may call a wavy line a "snake," (Gardner et al., 1978, pp. 16-18), or say "ball" to refer to peas, round beads on a necklace, pumpkins, pom-poms, and so on, or call various unfamiliar four-legged mammals "dog" (Anglin, 1983, p. 251). In these typical examples, the named analogous object is generally not present at the time the child is reminded of it by the stimulus that is present. This suggests that the analogies are based at least in part on the activation of long-term memories or knowledge representations of object shape.[3]

1.4. The Role of Shape in the Object Concepts of Adults

Most research investigating the role of shape in adult object classification has used recognition and classification tasks. In those tasks, participants must make rapid judgments about line drawings or photographic images of objects. These studies have shown that quick and accurate object recognition and classification is provided by overall object shape (Hayward, 1998; Rosch et al., 1976) and the shapes of object parts and their spatial interrelations (Biederman, 1987; Biederman & Cooper, 1991; Braunstein, Hoffman, & Saidpour, 1989; Tversky, 1989; Tversky & Hemenway, 1984; for a discussion, see Hoffman & Richards, 1984). Other object features, such as visible surface texture and colour, do not guide visual object recognition to the extent that shape information does (Biederman & Ju, 1988).[4] Shape was also found to be an important type of information when blindfolded participants identified objects through the sense of touch (Lederman & Klatzky, 1990). Other perceptual features can also be helpful for making finer categorical distinctions (Biederman & Ju, 1988; Lederman & Klatzky, 1990), but those fine distinctions often involve detailed analysis of shape information.

Shape analogies, of the type described above in young children's language, are reduced in adult discourse in proportion to other types of analogies, but they do continue to occur (Vosniadou & Ortony, 1983; Winner et al., 1980). In everyday discourse, terms like hammerhead shark, fiddlehead, needlenose pliers, and many others, are used widely and are understood easily, suggesting that shape remains an important dimension in adult object concepts. Poets and other writers may use shape analogies in expressing their ideas (e.g., Ammons, 1974). Scientists have reported shape-based analogies in their thinking. Perhaps one of the most famous anecdotes involved the chemist Kekule's mental image of a snake biting its own tail, thus forming the general shape of the hexagonal benzene ring (cited in Weisberg, 1986).

Shape-based analogies may be more common than we realize. For example, when we see a non-meaningful shape or pattern, sometimes we "see" a meaningful object that is similar in shape. People can be reminded of familiar things by stimuli such as clouds, ink blots (Thomas, Ross, & Freed, 1964), random polygons (Vanderplas & Garvin, 1959), or simple contoured shapes (Shepard & Cermak, 1973). In some studies, participants reported shape-based analogies spontaneously, though the researcher did not specifically request such responses (e.g., see Medin, Goldstone, & Gentner, 1993; Rosch, 1973). According to comprehensive reviews, the tendency to see one thing as another, sometimes called "seeing as" and often involving shape similarity, is common across societies, has appeared throughout history, and is evident in prehistoric artifacts and in the behaviour of numerous other animals (Guthrie, 1993; Janson, 1973).

A version of *seeing as* also can occur in thought, when no stimulus is present. This might be called "imaging as." For example, participants who mentally image an ambiguous figure, after having viewed it for only an instant, can sometimes image two or more different objects, one after another, that share the form of the original figure (for reviews, see Palmer, 1999; Peterson, 1993, 2003). People can, in their mental imagery, constructively combine simple geometric shapes, letters, and numerals to create or discover familiar objects and scenes (Anderson & Helstrup, 1993; Finke, Pinker, & Farah, 1989; Finke & Slayton, 1988). Using imagery, people can also combine known objects as parts in creatively constructing other known objects (Helstrup & Anderson, 1996). In completing these various tasks, people access known object or object part representations that are often similar in shape to the presented (or imaged) stimulus.

Simple geometric forms can remind us of numerous different objects. A triangle may be seen as a mountain, a wedge, an arrow (Wittgenstein, 1953/1973); a circle may be seen as a moon, eyeball, wheel, or pie (Barsalou, 1999; Gregory, 1973; also see the circles creativity test of Torrance, 1974). In a previous pilot study (Mattless & Anderson, 2000, unpublished), 53 undergraduate cognitive psychology students completed these "seeing as" tasks. Four stimulus figures for each participant were chosen from a pool that included line-drawn circles, triangles, rectangles, angles, and letters. For each of four different stimulus figures, participants typically were able to write the names of several everyday objects, within a two-minute time limit for each figure. For the circles, rectangles, and triangles, participants listed means of 12.1, 12.3, and 7.7 objects per shape, respectively. In other pilot work (Mattless, 2003, unpublished), each of two individually-tested participants were shown a simple line-drawn circle and were asked to say the names of as many objects (and/or object parts) as they could that had the approximate shape. Within eight minutes, one participant listed 53 objects and the other listed 85 objects or object parts. Consistent with the predictions of perceptual theories, objects that were more similar in shape (e.g., ball, grape) were reported clustered together in sequence more often than expected by chance, and were reported more quickly in succession than were differently-shaped objects (e.g., ring, pole).

1.5. Introduction Summary

Shape is a ubiquitous and important dimension in our experience with the world, and this is reflected in the organization of our knowledge and memory of objects. A similarity type of organization involving object shape appears to be established in our neural systems, as shown when people are reminded of an object by its shape similarity to another object. The fact that these organizational effects occur in everyday experience and in the experimental tasks considered thus far suggests that they may also occur in other tasks involving the free retrieval of object nouns. The next chapter will examine whether shape similarity organization is also evident in people's free emission and free recall performance.

1.6. Co-authorship Statement

The author (Paul Mattless) conducted all stages of the research, including the review of past research, the conception and development of the ideas regarding the shape similarity phenomena and the explanation of the possible underlying mechanisms, the conception and development of the ideas for the tasks and the experiments, the pilot studies and the experiments, the analyses of the data, and the writing of this thesis. The author's supervisor (Dr. Rita Anderson) contributed to this research though discussions, editorial comments, suggestions, and criticisms during all stages of the research and writing. The panel members (Dr. Mary Courage and Dr. Carolyn Harley) also contributed to this research through comments, criticisms, suggestions, and discussions, particularly on the proposal and the thesis text.

Chapter 2. The Role of Shape in the Organization of Object Memory: Free Emission and Free Recall

2.1. Clustering and IRT Phenomena in Free Emission and Free Recall

In standard categorical free emission (FE), also known as categorical fluency, the person's task is to list as many items as possible from a large category or domain (Bousfield & Sedgewick, 1944). For example, consider one person's retrieval sequence for the domain *animal*: "...mountain goat, sheep, chicken, duck, goose, mallard duck, turkey, porpoise, killer whale..." (Gruenewald & Lockhead, 1980, p. 231). This segment shows a typical pattern of clustering. A *cluster* is often defined as a unit consisting of two or more consecutive items that are similar or related in some respect. In this case, items are clustered into semantic categories that are subordinate to animal, categories that could be construed as four-legged hoofed mammals, fowl, and marine mammals. A temporal effect usually occurs with clustering: Interresponse times (IRTs) for transitions *within* a category, such as the transition between "duck" and "goose" in the above segment, tend to be shorter than those for transitions *between* categories, such as between "turkey" and "porpoise." I will refer to the within-category IRTs as *cluster* IRTs, and to between-category IRTs as *switch* IRTs.

In standard categorical free recall (FR), the study list is comprised of items from different categories or domains such as animals, names, professions, and vegetables (e.g., Bousfield, 1953; Bousfield & Cohen, 1953). The items are presented in randomized sequence, but participants, when recalling as many of the studied items as they can, tend to freely recall them in category clusters (for a review, see Kausler, 1974). The clustering and temporal effects that occur in FE also occur in FR retrieval sequences.

A critical aspect of standard FE and FR tasks is that participants are not given instructions or hints about how to produce, or study and recall, the items. Hence, any clustering that occurs is attributable to participants' own organizational tendencies as they attempt to report or recall as many of the items as possible (Kausler, 1974; Murphy, 1979; Murphy & Puff, 1982; Puff, 1979; Shuell, 1969; Tulving & Donaldson, 1972).

2.2. Contrasting Claims: Amodal Versus Perceptual Theories

The notion that thoughts and long-term memories are comprised of language-like amodal symbols has been a central assumption in many theories of cognition for the past few decades (for a review, see Barsalou, 1999, pp. 577-580). Over the past approximately 15 years, an increasing number of researchers have begun to challenge the amodal view, presenting arguments and evidence suggesting that knowledge is grounded in perceptuomotor and bodily representational systems (for overviews, see Barsalou, 2008; Barsalou, Simmons, Barbey, & Wilson, 2003). In the next section, I describe these competing theories in more detail, and define them within the context of this project.

2.2.1. Amodal Theories

According to some amodal theories (J. R. Anderson & Bower, 1980; Caramazza, Hillis, Rapp, & Romain, 1990; Pylyshyn, 1973; Shelton & Caramazza, 2001), perceptual information does not enter into long-term memories or knowledge. It is assumed that certain information from perceptual input systems, in the process of being encoded into long-term memory or knowledge, gets extracted and converted to word-like labels through an amodal symbol transduction process (Pylyshyn, 1984). The resulting symbols are said to be amodal because the format of the perceptual modality information has been removed. There is nothing visual and nothing spatial about an amodal symbol that represents a shape; it is merely a label that is in the same format as labels for all types of information. These amodal symbols are believed to be implemented in a central LTM system that is separate from the perceptual, motor, and specialized language systems in the brain. The word "apple" would have a lexical representation in the language module, plus an amodal label for the concept of apple in the central amodal LTM system.

All amodal theories, as defined here, postulate that categories, relations, and concepts in LTM are represented in a central amodal semantic LTM system by arbitrary, language-like amodal symbols (i.e., "Mentalese", see Fodor, 1983). In this critical respect, even an otherwise multimodal theory of memory, such as that of Engelkamp and Zimmer (1994), is amodal. Although their theory is multimodal in that it retains a perceptual format with respect to item-specific perceptual information for individual objects, it is amodal with respect to categorical and relational representation.

I focus on Engelkamp's and Zimmer's (1994) theory because it is one of the most advanced and explicit with respect to memory organization of objects and human actions. They state clearly that object nouns having similarly shaped referents should not be clustered in recall. In their view, categorical clustering in recall depends on the activation of amodal categorical labels that link to other amodal labels for members of their respective categories. Activation of modality-specific item-specific shape information, as might occur when a person mentally images each object, should have no effect on recall organization. Engelkamp and Zimmer (1994) believe that amodal labels for shapes may already be established in a small number of instances where there are strong preexisting associations between the object name and the shape name (e.g., between *ball* and *round*), or may be formed anew when the shape similarities among the stimuli are made very salient due to task instructions, cuing, or obvious stimulus characteristics. However, they also maintain that object shape category is not normally represented amodally for most objects and is not normally converted to amodal labels in FR of nouns (Engelkamp & Zimmer, 1994; H. D. Zimmer, personal communication, November, 2001). If they are correct, there should be no significant clustering by noun referent shape where (a) the word association strengths between the object noun and the shape label are minimal, (b) the shapes are not overly salient, and (c) the participants are not cued, told, oriented, or given hints to group items by shape. Thus, in standard FR tasks, *same or similarly-shaped items should not be clustered significantly in recall.* Given the dependency between clustering and IRT effects, a logical implication is that *shape cluster IRTs should not tend to differ in duration from shape switch IRTs.* Logically, these hypotheses can be extended to clustering and IRT effects in FE. This claim is also consistent with Shelton's and Caramazza's (2001) assertion that, in tasks such as standard lexical decision, there should be no response time savings in judgements of perceptually-similar (e.g., coin-pizza) versus perceptually-dissimilar (e.g., coin-paint brush) prime-target noun referents (e.g., see Pecher, Zeelenberg, & Raaijmakers, 1998).

2.2.2. Perceptual Theories

Some perceptual theories posit that objects are represented in an LTM system that has uni-, multi-, and supramodal[5] bases (Barsalou, 1999, 2008; Barsalou et al., 2003; Fuster, 1995; Hebb, 1949, 1968; Martin, Ungerleider, & Haxby, 2000). In these views, information from multiple early unimodal sources is coordinated and integrated in later unimodal, multimodal, and supramodal neural assemblies that are distributed widely throughout numerous regions of the brain (e.g., see Fuster, 1995), not by an amodal system. The word "perceptual" is not limited to its usual meaning, namely, perception of stimuli in the environment,[6] but also refers to conceptual, emotional, and memorial representations. According to perceptual theories, remembering or thinking of an entity, in its absence, involves reactivation of a subset of the neural systems that were active in previous experience with that entity. A sampling of the reactivated experiential information could include the physical aspects of the entity. For example, if we think of a snake and its long flexible form, movement, and other physical aspects, perceptual and motor areas of the brain that process form and movement are activated. Although these aspects of conceptual information may become associated with word representations, or may often be represented in an abbreviated and schematic manner, they retain a perceptual format; they are neither converted to nor replaced by amodal symbols.

Consistent with perceptual theories, I propose that shape categorical representations in LTM are implemented in a subset of the neural populations that are active when we perceive and interact with objects (Barsalou, 1999; Hebb, 1949, 1968). In extending perceptual theories, I predict that when retrieving a noun (e.g., snake) in a FE or FR task, some perceptual properties of the referent will tend to be reactivated, including its form. I posit that the neural assembly for a snake's form shares form-selective neurons with those of other entities' forms (e.g., rope, hose, etc.). Once activated, these neurons may stay active, or may become sensitized temporarily, thereby increasing the chances of activating other object representations that include the same form-selective neurons. In part because of this overlapping implementation between similar object representations, in FE and FR, *similarly-shaped items should tend to cluster significantly, and shape cluster IRTs should tend to be shorter than shape switch*

IRTs. Alternative clustering schemes that inevitably arise in FE and FR should involve other perceptual similarities, spatial and action relations, and so on.

2.3. Does Clustering Occur for Perceptual Categories and Relations?

The clustering of words by semantic categories and relations in FE and FR retrieval is one of the most well-established of phenomena in all of memory research. Several types of semantic conceptual clustering have been found in the retrieval sequences, including associative (Deese, 1959; Jenkins & Russell, 1952; Pollio, 1964; Pollio, Kasschau, & DeNise, 1968), categorical (Bousfield & Cohen, 1953; Bousfield & Sedgewick, 1944; Cohen, 1963), schematic (Khan & Paivio, 1988; Rabinowitz & Mandler, 1983), script (Bower & Clark-Meyers, 1980), action (Koriat & Pearlman-Avnion, 2003; Zimmer & Engelkamp, 1989), musical genre (Booth & Cutietta, 1991; Halpern, 1984), and social contextual (Bond, Jones, & Weintraub, 1985; Fiske, 1995). The usual speed advantage for cluster over switch IRTs has also been found in a variety of categories in FE (Bond et al., 1985; Graesser & Mandler, 1978; Gruenewald & Lockhead, 1980; Rosen & Engle, 1997; Rubin & Olson, 1980; Yumino, 1977) and in FR (Ashcraft, Kellas, & Needham, 1975; Bjorklund, 1988; Hasselhorn, 1992; Kobasigawa & Orr, 1973; Pollio, Richards, & Lucas, 1969; Patterson, Meltzer, & Mandler, 1971; Rubin & Olson, 1980; Wingfield, Lindfield, & Kahana, 1998). There have been some failures to demonstrate semantic clustering (see Shuell's 1969 review; also see Dabady, Bell, & Kihlstrom, 1999), but such failures are uncommon for most semantic categories tested.

Despite the generality of semantic clustering and IRT effects, several studies, to be reviewed briefly in this paper, have found that clustering of items from perceptual categories is not much more (if more) than would be expected on the basis of chance.

This suggests a discrepancy, where perceptual clustering is seen as an exception to the general trend of clustering for non-perceptual similarities (categories) and relations. At first glance, the apparent lack of perceptual clustering seems to support the conclusion drawn by some researchers that perceptual categories, unlike semantic categories and associations, are not represented in LTM and play little or no role in the organization of recall (Engelkamp & Zimmer, 1994; Perrig & Hofer, 1989).

Is it really the case that perceptual category clustering has not been demonstrated? Have some instances of perceptual clustering been overlooked? Is semantic clustering of concrete entities influenced by perceptual similarities? Evidence will be brought to light that raises doubts about the presumed non-occurrence of perceptual clustering. I will explore the most likely reasons why some kinds of perceptual clustering have not been shown clearly. I will then present new empirical evidence bearing on this issue, particularly in regards to shape categorical clustering.

2.3.1. Evidence of Primary Perceptuomotor Clustering

Primary clustering occurs in FR or FE when items are clustered according to characteristics that are physically present in the stimuli or responses. For example, in the first-letter or first-phoneme fluency (FE) task, where the goal is to list as many words as possible that start with the same letter, clustering occurs according to sounds or letters that overlap between the successive words. If the task is to list words that start with *s*, a person might list "...spot, spill, subject, swish, swell, switch, spell..." (cited in Kolb & Whishaw, 1990, p. 475; also see Troyer, Moscovitch, Winocur, Alexander, & Stuss, 1998). Words that rhyme may also be clustered in FR (Bousfield & Wicklund, 1969; Fagan, 1969). Clustering by perceptuomotor similarities or relations has been found in FR drawing of abstract geometric figures (Bousfield, Berkowitz, & Whitmarsh, 1959)[7], abstract geometric figure drawing fluency/FE (Jones-Gotman & Milner, 1977)[8], FR drawing of places or landmarks on a map (Taylor & Tversky, 1992), gesture fluency/FE (Jason, 1985)[8], FR of melodies (Cutietta & Booth, 1996), and FR of actions (Koriat, Pearlman-Avnion, & Ben-Zur, 1998; Koriat & Pearlman-Avnion, 2003; but for null clustering results see Zimmer & Engelkamp, 1989). Primary clustering evidence demonstrates that there is nothing inherent in the nature of perceptual representations or systems that precludes categorical or similarity-based organization.

2.3.2. Evidence of Secondary Perceptual Clustering

Secondary clustering occurs in FR or FE when items are clustered according to similarities or relations that are not present directly in the physical or temporal aspects of the stimuli and responses, but which occur in the referents. Semantic category clustering in the retrieval of nouns is an example of secondary clustering, but a perceptual version of it is also possible. In a FR pilot study (n = 6; Mattless, 2003, unpublished) where object names were presented aurally in random order, participants showed significant clustering by shape in their spoken retrieval output (e.g., "...rope, snake, noodle, hose, pearl, bubble, globe..."). Likewise, shape cluster IRTs were significantly faster than shape switch IRTs. Although some previous FR studies reported no significant clustering of object nouns by shape categories (e.g., Zimmer, 1989), those studies, as will be shown later, contained some methodological limitations that preclude clear interpretations of the results.

Entities that are close together spatially in the known environment also tend to be clustered together in retrieval output. This has been shown in FE tasks, such as in the retrieval of the names of home furniture objects (Plumert, 1994), campus buildings

(Hirtle & Jonides 1985), cities (Bousfield & Sedgewick, 1944), and states (Shepard, 1957). These results support the assumptions of perceptual theories (Barsalou, 1999; also see Brewer, 1999). Although Engelkamp and Zimmer (1994) claim that clustering of known objects in their known relative locations is based on amodal representations, they had further implied that spatial relational clustering does not normally occur under standard FR when the to-be-recalled items are randomly-positioned and unrelated. However, again in support of perceptual theories, there is considerable evidence that when unrelated items are presented in randomized spatial positions at study, they are later clustered by inter-item spatial proximity in standard written or oral FR (McNamara, Hardy, & Hirtle, 1989; Moar, 1977; Nida & Lange, 1997; Plumert, 1994; Plumert & Strahan, 1997; Stukuls, 1975). There was also evidence of secondary spatial relational organization in the pilot FR study cited above (Mattless, 2003, unpublished), where some participants reported that they had organized some unrelated objects according to spatial relational schemes.

2.3.3. Perceptual Similarities May Contribute to Semantic Clustering

Multiple perceptual similarities and relations between items may contribute to clustering in the retrieval of items from concrete semantic categories. This is not merely possible but is likely because semantically similar objects/entities tend to be similar and related in numerous respects, including global shape, shapes of parts, presence of specific parts (Tversky & Hemenway, 1984), size (Caramazza, Hersh, & Torgerson, 1976; Henley, 1969), schematic structure, surface properties, various other perceptual properties; and in associated locations, actions, uses, affective states, and so on (Damasio, 1990; Humphreys & Forde, 2001; McRae, de Sa, & Seidenberg, 1997; McRae &

Boisvert, 1998; McRae, Cree, Westmacott, & de Sa, 1999; Rosch & Mervis, 1975; Rosch et al., 1976; Small, Hart, Nguyen, & Gordon, 1995). Inspection of concrete semantic FE output sequences, such as of animal names (e.g., as in the example from Gruenewald & Lockhead, 1980, cited previously), supports this hypothesis. Many past FR studies that used semantic category lists included names of objects/entities that were physically-similar. For example, Bousfield's (1953) *animals* category included, among others, "woodchuck, otter, weasel, badger" (p. 230). In FR, when the concrete noun referents within semantic categories are high in feature overlap, there is greater clustering than when there is low feature overlap (Cissè & Heth, 1989), and many of these features are perceptual properties. In other tasks, when perceptual or motor information pertaining to concrete entities is retrieved from LTM, there is activation in appropriate perceptual and motor cortical areas (for reviews, see Martin, 2001, 2007; Martin et al., 2000; Ungerleider, 1995). All of this indirect evidence supports the idea that concrete semantic category clustering depends partly on the activation of overlapping perceptual features, plus numerous other similarities and relations.

2.3.4. Problems in Studies Showing Low Perceptual Clustering

Unlike the concrete semantic clustering examined thus far, which involves the confounding of many similarities and relations within categories, the items within perceptual categories used for FR lists have usually contained only a single shared feature (e.g., a specific colour) between exemplars. In this respect, the standards for demonstrating secondary perceptual category clustering have been more stringent than the standards for demonstrating semantic category clustering. To complicate matters, there were numerous major methodological problems in past perceptual category clustering research, including the use of items that were not very similar within categories. Yet, to demonstrate perceptual category clustering experimentally, the items within the categories should not be similar in other respects. Quite rightly, if perceptual clustering according to one set of target categories is to be demonstrated, it must involve switching across multiple non-target perceptual similarities, semantic categories, and other interitem relations. For example, clustering based almost entirely on shape similarity would include a sequence such as "...volleyball, apple, pearl..." where there is switching between sports, foods, and jewellery categories. Hence, perceptual theorists should expect that mean levels of clustering for shape will be lower than for concrete semantic categories, but significantly above chance, if the methodological problems can be sufficiently reduced.

2.3.5. Summary

The existence of various kinds of perceptual clustering raises questions about the assumptions in some amodal theories. Some concrete semantic clustering in FE and FR may incorporate perceptual similarities between exemplars. The lack of significant secondary perceptual categorical clustering in some previous FR studies could be due to various methodological limitations [discussed further in Study 2: Free Recall (FR) Review]. It is reasonable to suppose that perceptual organization effects could occur if the methodological problems were reduced. No previous FE studies have examined secondary perceptual categorical organization effects. Hence, in the next section, I examine whether perceptual clustering and temporal effects can be observed in the FE of object nouns.

2.4. Study 1: Free Emission (FE)

To test the perceptual versus amodal hypotheses, this study set out to examine clustering and IRT evidence from FE tasks. In these tasks, people listed objects from superordinate shape domains. These domains were *long*, *flat*, *round*, and *straight-sided*. Each individually-tested participant listed objects continually for 7 minutes within one domain (e.g., listing *long* objects), and then, after a distracter task interval, listed objects for another domain (e.g., *flat* objects) for 7 minutes. Within each domain, there are different three-dimensional shape subcategories. For example, within the shape domain *round*, objects listed may be spheroids, rings, disks, and cylinders (e.g., grapefruit, hula hoop, plate, and pop can, respectively). The shape domain tasks conform to the standard categorical FE procedure: A verbal label of the superordinate category is presented; the possible subcategories are not presented; no information is given regarding clustering schemes; and participants naturally may use a variety of other clustering schemes (spatial contextual, semantic categorical, etc.). Therefore, the subcategorical shape clustering that could be found in the tasks here should be considered as valid an indication of organization as that found in other FE tasks.

2.4.1. Why Focus on Simple Shapes?

There are several good reasons for focussing on simple shapes. *First*, the view that objects are not organized by shape category in LTM (Engelkamp & Zimmer, 1994) was based primarily on Zimmer's (1989) FR results that suggested a lack of shape clustering. His study lists contained object items from simple shape categories such as spherical, triangular, stick-like cylindrical, flat rectangular, cube/block-like, ring-like, disks, string-like, and board-like. These shape categories can be sampled in FE under round, straight-sided, flat, and long object domains, making this study comparable to Zimmer's (1989) FR study with respect to categorical content. Many of the object names produced from these simple shape FE tasks will later be used in the FR study (Study 2).

Second, thus far, the role of three-dimensional shape in object representation has generally been investigated using tasks that present drawings or photos of objects (for discussions, see Biederman, 1987; Connor, 2004; Peterson & Rhodes, 2003). No previous FE studies have focussed explicitly on the role of shapes in the organization of object memory, when no real or depicted object or shape information is presented.

Third, although Rosch et al. (1976) demonstrated the importance of shape in people's basic-level object categories (e.g., cat, bicycle, etc.), many of their object shapes were relatively complex. Simple shapes, compared to complex shapes, tend to be less confounded with semantic category membership. For example, simple spheroid objects can occur in a variety of different semantic categories, such as sports and games, mechanical parts, and fruits and vegetables, but dog-shaped things (for lack of a better term) are generally only dogs, related species, or artificial representations of dogs. Because there is somewhat less confounding of shape and semantic category for simple-shaped objects, the FE shape domains used here allow some opportunities for sampling similarly-shaped objects whose pairings in the output series cut across semantic categories. In pilot work, other FE domains were tried, such as *objects*, *things that can be held in one hand*, and *containers*. The shapes of the objects reported from those other domains tended to be more complex, and more confounded with semantic category, than the predominantly simple objects of interest here.

Fourth, simple object shapes should be more easily classified and rated by judges than complex shapes. The shapes of parts play an important role in object classification for more complex objects (e.g., see Tversky & Hemenway, 1984, who re-analyzed some of the data from Rosch et al., 1976), though in the context of the present study would probably make the classifications and ratings more difficult. Relative to other FE domains (animals, objects, etc.), those used in the present work seemed to reduce the percentage of complex objects reported, making the results easier to interpret.

Finally, simple object forms occur in large numbers in our everyday experience. This is shown not only by observation, but had been revealed in people's FE output in previous pilot work (Mattless, 2003, unpublished).

2.4.2. Overview of the FE Study

The FE study had three phases. In Phase 1, participants reported as many objects as they could think of within a general shape domain. Each individually-tested participant completed FE for two different domains, one domain at a time. After participants completed the second of the two FE tasks, they described their thought processes that occurred during the FE tasks. They then classified and rated the objects in their transcribed output sequences according to shape category and shape similarity.

Phase 2 involved preparing the materials produced in Phase 1 for measurement and scoring. Object names produced, classified, and rated by FE participants in Phase 1 were also classified and rated by other participants in Phase 2. For the set of items from each shape domain, different participants sorted cards bearing the object names according to three-dimensional shape, conceptual semantic, or ordinal size categories. Other participants rated the shape similarity between adjacent items in the FE output lists. Interjudge agreement was assessed for the classifications and ratings.

In Phase 3, the researcher attempted to obtain shape dominance and typicality scores for object names, for possible use in the selection of study list materials for the subsequent FR study. For these tasks, participants (a) listed the names of objects for numerous specific shape categories that were identified in the Phase 2 classification procedure, and then (b) made shape typicality ratings on large numbers of objects primarily from a selection of the Phase 1 FE participants' output. Results for the object listing and typicality rating tasks are presented in Appendix A. Only the typicality ratings were used in selecting the items for the subsequent FR study.

2.4.3. Predictions

Four shape domains (long, flat, round, straight-sided) were used so that a wide variety of simple object shapes would be sampled. Comparisons between the conditions were not of central interest. This FE study was designed primarily to determine whether or not there is evidence for the existence of shape organization of objects in human memory, not to compare domains.

To the extent that semantic and shape categories are naturally confounded in most objects, including among those having simple shapes, there should be a positive association between people's shape and semantic clustering scores. There should be some overlap between semantic and shape classifications when judges classify the sets of object items. One should also reasonably expect some confounding between object size and semantic category (e.g., food items should tend to be smaller than furniture items). Both shape and semantic categorical clustering were expected to reach significance, with semantic clustering generally higher than shape clustering.

According to perceptual and amodal theories, semantic cluster IRTs should tend to be shorter than semantic switch IRTs, but only the perceptual theories predict that shape cluster IRTs should be shorter that shape switch IRTs. The perceptual but not the amodal theories also predict that as the shape similarity between sequential pairings of object items increases, IRT durations should be reduced, even when other factors such as semantic IRT type (cluster versus switch) are partialled out statistically. In addition, only perceptual theories predict that as the size difference between listed object pairings increases, IRT durations should also increase, even with other factors partialled out.

Perceptual, but not amodal, theories predict that when people are requested to access knowledge of concrete entities, they construct experiential or perceptual simulations of the entities and the contexts in which they occur. People should report that they used imagery to access object items, even though there are no instructions or hints to use imagery.

2.5. Methods for the FE Study

2.5.1. Participants

In the first phase of the study, 22 Memorial University students participated in the FE tasks (11 females, 11 males; median age = 18.5). The majority were enrolled in a first-year psychology course. Eight of the 22 were pilot participants.

In the second and third phases of the study, 46 Memorial University students participated in the various classification, rating, and listing procedures (23 females, 23 males; median age = 19). Seven of these participants (4 females, 3 males) completed pilot versions of the classification and rating tasks.

The research project, including all aspects of the FE and FR studies, was approved by the Interdisciplinary Committee on Ethics in Human Research of Memorial University and was carried out in accordance with the Tri-Council Policy Statement on Ethical Conduct for Research Involving Humans. The author (Paul Mattless) was the experimenter throughout the research project. All participants were volunteers and were paid $6.50/hr, which was the standard rate set by the Psychology Department. A typical session for an FE study participant was 2 hrs long, but some were 1 hr 30 min long.

2.5.2. Procedures for FE Phase 1: FE and Post-FE Tasks

2.5.2.1. Prior Knowledge of the Study and Task Content. During the informed consent procedure at the start of each session, each participant indicated to the experimenter's satisfaction that she/he did not have specific prior knowledge of this study or the tasks used in it, and had not participated previously in any word-listing or recall tasks. At the end of the session, at the end of debriefing, each participant was asked not to tell other students specific information about the research hypotheses or about the shapes and shape categories. The importance of this issue was explained to the participant.

2.5.2.2. The Testing Room and the Recording Apparatus. To minimize visual contextual cuing, testing occurred in a small room with no windows and few visual distractions. The room contained a long table, two chairs, a small garbage can, a small recycling can, and a closed cabinet. There were no windows or mirrors. The walls were bare. Only the experimenter and the participant were in the room. They were both seated on the same side of the table. The experimenter and participant usually each had a knapsack and jacket in the room, placed behind them on a cabinet top and door-mounted coat hanger. A limited number of objects necessary for the procedure were visible in the room. On the table there were papers, a pen, pencil, eraser, stopwatch, digital audio recorder, audio tape recorder, a wire from each of the recorders to the outlets, a small earphone (for the experimenter), and a clip-on microphone (to be worn by the participant) with a wire to one of the recorders. Occasionally, a participant happened to bring an additional object (e.g., a cup of coffee, or a binder, or some such school-related item) into the room. For those few cases, the experimenter decided to not draw special attention to the object and therefore did not ask that it be removed.

The spoken retrieval output was recorded using a Sony ICD-ST25 digital voice recorder and a Sony TCM-939 audio cassette tape recorder. The participant wore a Sony ECM-T6 tie-clip condenser microphone connected to the cassette tape recorder.

2.5.2.3. FE Conditions. Each individually-tested person completed two FE shape domains, one domain at a time. To minimize the number of potential objects that could be listed for both domains by the same participant, participants in one group received long and flat, and those in the other group received round and straight-sided. The order in which the domains were tested was counterbalanced within each of those groups.[9] Each person was assigned randomly to one of the two groups, but with the restrictions that there be near-equal numbers of people, and of males and females, per group.

Each individually-tested participant completed several tasks, all within a single two-hour session, in the following sequence:

1. Practice Domain FE, Objects
2. First Shape Domain, FE
3. Arithmetic Distracter Task
4. Second Shape Domain, FE
5. Questions About FE Thought Processes
6. Object Shape Classification for First Domain
7. Shape Similarity Rating for First Domain
8. Object Shape Classification for Second Domain
9. Shape Similarity Rating for Second Domain.

2.5.2.4. Practice FE Task Instructions. The practice task was constructed to prepare participants to list real three-dimensional objects in the shape domain tasks. All except 3 pilot participants[10] in this study first completed a very brief (duration = 10 s) practice task, listing as many real physical things (objects) as they could. The experimenter avoided providing any examples during the instructions, but did give a clear brief answer (e.g., "yes," "right"; or "no") if a participant provided an example in asking a question. The instructions were lengthy because, without the specifications, pilot results suggested that participants would list some items that were not real three-dimensional objects having definite boundaries, or at times would include irrelevant speech. The practice task itself was made very brief (10 s) to prevent participants from having the time to practice extensively any organizational schemes prior to the shape domain FE tasks. No mention of shape was made until the shape domain instructions were given. Instructions for the practice task were as follows:

> The audio recorders will be turned on and you'll have your eyes closed during this brief practice task. The practice task will be to list the names of real physical things for 10 seconds. A real physical thing is any individual, whole, physical body or object that exists. I want to emphasize some parts of this definition, because it's important to follow these guidelines when you are thinking of objects or things and reporting them. *Individual* means a single thing or object; those are what you should think of and report. Say single object names; there should be no *s* on the end of any word. *Whole* means whole things, not parts or aspects. *Physical body or object* means real physical things or objects that exist, but does not include things that are just substances and surfaces. Are you clear on what kinds of things to list? You will need to keep these criteria in mind later when you are doing the longer listing tasks.
>
> When you list the names, just say the name of each thing, one after another, and nothing else. You don't have to describe or justify the things you name. We can talk about certain items, if necessary, after the task. During the task, say only the names of the things, as soon as you think of each one. I'll let you know when the time is up. No questions can be asked once you start, so if you have any remaining questions, you can ask them now. (Pause). If you're ready, please close your eyes and (Starting recorders here) say the names of as many real physical things as you can think of, starting now.

2.5.2.5. Shape Domain FE Instructions. FE instructions for a participant's first domain trial (e.g., flat) were as follows:

> This next task follows the same format as the practice, except it's longer and more challenging. This task will be to list, for seven minutes, the names of as many different flat things as you can think of. (Definition) A *flat* thing is any individual, whole, real thing that exists as a distinct physical body or object that is relatively flat, meaning simply that it is wide or broad, but not very thick. You probably know hundreds of flat things. The task now will be to report as many of them as you can. Production is the main goal: You want to list the maximum number of flat things that you possibly can. Also, the flat things that you list should indeed be relatively wide or broad, but not very thick.
>
> I can tell you that thinking of more and more flat things will get tougher after the first few minutes. That's normal. It's important to keep your mind focussed on the task; this will help you continue to come up with flat things throughout the seven minutes. Again, you'll have your eyes closed during the task, and you'll open them when the time is up. I'll let you know when the time is up. During the task, say only the names of the things, and nothing else. Are you ready? (Pause for any questions or comments). Close your eyes and (starting recorders here) say the names of as many different flat objects as you can think of, starting now.

FE instructions for the other domains followed the same format for *flat*, above. The distinctive elements for the *long* domain were as follows:

> …any…thing…not very broad, and not very thick in proportion to its length…could also be described as slim, or slender, or elongated…(long) objects should typically be at least five times as long as any of their other respective dimensions, such as width or thickness.

Without those stipulations, pilot results suggested that participants would mention some objects that were not particularly long, or not likely to be described as such. The stipulation also reduces overlap between long objects and the other domains, especially flat. Both the flat and long definitions were based partially on dictionary definitions. The distinctive parts of the instructions for round and straight-sided domains were as follows: "(round)…any…thing…round."

"(straight-sided)...any...thing...whose sides are all at least roughly straight."

FE instructions for the second domain were similar to the format of the first, but were briefer for parts that were already described in the first.

2.5.2.6. Arithmetic Distracter Interval Task. A distracter interval task was used to give participants a rest from listing objects and to reduce interference from the first FE task performance on the second FE task. During the 3-minute interval between the first and second FE tasks, the participant completed pencil-and-paper arithmetic problems. Half the participants did addition and the other half did subtraction. Participants were instructed to work continuously and accurately in completing as many of the problems as they could.

2.5.2.7. Questions About FE Thought Processes. After the second shape domain FE task had been completed, participants answered an open-ended question about the thought processes they had used in each of the FE tasks. In total, answering the two questions took approximately 5 minutes. The participant wrote an answer to the written question (below) for one shape domain, and then answered the same question as it applied to the other shape domain.

> 1(a). Describe the ways in which were you able to think of so many (e.g., flat) things. You may use specific examples in your descriptions. 1(b). If applicable, also describe any difficulties you may have had in coming up with more (e.g., flat) things. (If you need more space for your descriptions, please continue on the back of this sheet).

All participants answered the questions. No one reported any difficulties in remembering thought processes that were experienced during the FE tasks.

The written responses to post-task questions regarding thought processes experienced during FE were classified according to participants' descriptions. This classification scheme is described later in the Results section, 2.6.7., see Table 21.

2.5.2.8. Transcription of the FE Output Sequences. Listening to a digital recording on earphone (not audible to the participant), the experimenter transcribed the first domain output items onto an 8.5 x 14" sheet while the participant was working on the arithmetic and the post-task questions. The recording of the second domain was transcribed while the participant worked on the grouping and rating tasks for the first domain output transcript, tasks that are described next.

2.5.2.9. Object Shape Classification. The purpose of this task was to obtain the shape classification appropriate for each listed object. During transcription, the experimenter put a question mark beside any listed object names on the transcript that did not clearly imply one specific kind of three-dimensional shape (e.g., listed for *round*, "candy" could be disk-like, spherical, cylindrical, etc). In the first step of object classification, for those items having a question mark, the participant was instructed to do a small rough outline sketch of the specific object and write a more specific name of the exemplar of the object that they remembered having listed. If the exemplar could not be remembered, the participant was instructed to make another question mark and then draw and name the exemplar most familiar to him/her (for the detailed instructions, see Appendix B). When asked, participants were almost always quite certain of their classifications about the exemplar object shape that was thought of and listed during FE.

Next, the task was to classify all of the objects on the transcript by shape. The participant was instructed to write a letter (any letter from the alphabet) standing for the three-dimensional object shape category for each object name on the transcript. The participant made a drawing of the shape category for an example object in each category (for detailed instructions, see Appendix B). The experimenter used an example transcript sheet to illustrate the procedure, though no actual object names or shapes of interest were shown. In this process, the participant supplied the shape categories and assigned objects to those shape categories. The experimenter examined the classifications when they were complete. Classifications were discussed when there was a need for clarification, or when the participant had a question, or when there was an apparent error or inconsistency. In all instances, the participant made the decision for the shape category assignment of the item.

2.5.2.10. FE Shape Similarity Ratings. The purpose of this task was to obtain a shape similarity rating for each successive pairings of items. In this task, using the output transcript, the participant was instructed to rate, on a scale of 1–7, the three-dimensional shape similarity between the successive adjacent pairings of list items. The shape similarity scale was as follows: 1 = very low; 2 = low; 3 = moderately low; 4 = moderate/average; 5 = moderately high; 6 = high; 7 = very high. The participant rated the similarity between the first and second objects, second and third, third and fourth, and so on. The experimenter examined the ratings when they were complete, and pointed out any apparent errors or internal inconsistencies (e.g., if a transition between a spherical item and a disk-like item was given a "2" in one case but a "4" in a different case). Regarding those editorial comments, the participant was informed that she/he had the final decision on the rating. For detailed instructions, see Appendix B.

2.5.2.11. Second Domain Shape Classification and Similarity Rating. Object Shape Classification and Shape Similarity Ratings were obtained using the same methods and in the same order as in the previous section. Instructions were briefer than those for the first domain. Each participant was debriefed at the end of the experimental session.

2.5.3. Procedures for FE Phase 2: Classification and Rating of the Materials

In Phase 2, new participants classified object items according to shape, ordinal size, and semantic category. Another group of participants completed shape similarity ratings of the previous FE participants' object lists. All of the participants in Phase 2 were tested individually in the same room setting as those in Phase 1. Again, the experimenter asked the participant not to discuss with other students the hypotheses or contents of the tasks, especially with regard to shapes and categories.

For the common everyday object items produced in this experiment, shape and size classifications and ratings from previous norms were not available. Previously, Shepard and Chipman (1970) had participants in one of their conditions rank the shape similarity of pairings of 15 of the U.S. states based on memory, with no depictive state shape stimuli presented during that condition. To my knowledge, no psychological researchers have previously had people classify large numbers of everyday objects according to shape, using only the object names. In addition, available classifications or ratings for object size contained too few items to be of use for the present samples. Semantic category norms for object items exist, but to my knowledge, at the time of testing, they were out of date and did not contain many of the items from the present samples.

Interjudge agreement was assessed for each of shape and semantic classifications, size ranking, and for shape similarity ratings. Levels of agreement for size ranking and semantic category, respectively, were expected to be high, but the reliability of people's object shape classifications had not been tested previously.

2.5.3.1. Shape Similarity Ratings (Judges). Seven individually-tested judges (4 males, 3 females) carried out shape similarity ratings on copies of the output transcripts, following the general procedure described for the original FE participants' shape similarity ratings (see Appendix B). One judge completed two rating sessions, and two others had previously completed the full FE procedure. Other than those cases, the shape similarity judges did not complete any other task in this study. Judges were informed that they would be rating lists produced by other participants. They were instructed not to rate any pairings that contained an item that they did not know. Each output transcript copy contained all the items produced, but did not show the classifications and ratings by the original FE participant. Specific drawings or descriptions for ambiguous items were listed or copied from the original FE participant's transcript. Judges were able to complete a mean of seven FE transcripts within an approximately two-hour session. For each FE transcript, there were shape similarity ratings from two judges ("Judge 1" and "Judge 2").

2.5.3.2. Shape Classification of the Master Sets. All the different object names produced within each domain, pooled from FE participants' output for that domain, were transcribed onto 5 x 9 cm cards, with one object name per card. In some cases, for items originally identified as ambiguous (e.g., "candy"), additional descriptive and depictive information provided by the original FE participant was noted or copied onto the card. The numbers of *different* items per domain were as follows: Long = 259; Flat = 245; Round = 302; Straight-Sided = 260. Each of these four decks of cards constituted the master set for its respective domain. Cards were shuffled thoroughly before every individual session.

Each of fourteen individually-tested participant-judges sorted the cards for a domain's master set into separate groups according to the overall three-dimensional shape similarities and differences between the objects. There were 7 females and 7 males, assigned to conditions randomly under the restriction that, for each shape domain, there were at least three judges, and at least one male and one female judge. None of the shape judges performed any other tasks in this study. The experimenter instructed the judge to take into account multiple shape features such as curvature versus straightness, proportion of length to width, relative thickness, shape flexibility, and structural shape complexities of the objects. The judge was also instructed to ignore other aspects like size, orientation, material composition, and conceptual/semantic category (e.g., food, sports item, etc.). Each judge completed the classification of a single shape domain, taking approximately two hours. During the last approximately 20 minutes of the two-hour session, the experimenter checked over and obtained explanations and descriptions from the judge for his/her groupings. Detailed instructions for this task can be found in Appendix C. Once the sort was complete, the arrangement of the items was transcribed by the experimenter for subsequent analyses.

2.5.3.3. Shape Classification Measurement and Scoring. Under the shape classification procedures, participants formed their own categories of the object items. Therefore, the Brennan and Light (1974) approach was appropriate to measure the level of inter-judge agreement when two independent judges make their own categories, in classifying the same set of discrete items. In this scoring scheme, an agreement occurs each time a given pairing of objects, such as *a* and *b*, is grouped together by Judge C [e.g., in the group (*a, h, b, k*)] and is grouped together by Judge R [e.g., in the group (*f, b, e, g, a*)]. That is a same-same agreement. A second type of agreement occurs when the items in a given pair of objects are each assigned to separate groups by Judge C and Judge R (i.e., both judges assign *a* to a different group than *b*). This is a different-different agreement. Disagreements occur when Judge C puts *a* and *b* in the same group but Judge R puts *a* and *b* in different groups, or vice versa. Schematically, Judge R's categories represent the rows, and Judge C's categories represent the columns, of a rows x columns table. The agreements are then entered into the appropriate cells of the rows x columns table in preparation for analysis. As recommended by Brook and Stirling (1984), ordinary χ^2 tests of association between judges were obtained.[11]

2.5.3.4. Semantic Classification of the Master Sets. Eight individually-tested participant-judges (4 females, 4 males; assigned randomly but with the restriction that there be 1 male and 1 female per domain) each sorted the cards from a single domain into groups according to conceptual categories and relationships, using the same procedures as for the master sets shape classification task, described above. The judge was instructed to focus on the general conceptual category of the objects (e.g., sports-related, fruit or foods, kitchen utensil, etc.), and to ignore physical aspects such as shape, size, or material. Judges typically completed the procedure in about 90 minutes. None of them performed any other task in this study. Detailed instructions for this task can be found in Appendix D. The same measure of agreement as described for shape was used for finding the levels of semantic category agreement for each pair of judges.

2.5.3.5. Composite Classifications for Shape and Semantic Master Sets. A single "average" shape classification scheme was constructed by the experimenter based on the judges' master set classifications. This single composite was constructed for the shape classifications so that category assignments of items in FE participants' output lists could be determined against a standard scheme. A semantic composite was also constructed for that purpose. To verify that the composite actually reflected the judges' classifications, agreement was also measured between each judge and the composite (see below). The composite classification for each domain is presented in Appendix N.

2.5.3.6. Size Classification of the Master Sets. Four individually-tested participant-judges (2 females, 2 males) each assigned the cards to the ranked columns according to the relative size of the objects. Each judge completed two domains, one domain at a time. One male and one female ranked long and flat domains, and the other male and female ranked round and straight-sided domains. None of the size judges completed any other task in this study. The experimenter set up headings numbered from 1 to 9, making 9 ordinally ranked columns, where 1 represented the smallest objects and 9 represented the largest. In placing the cards into the columns, the judge was instructed to take into account the length, surface area, and volume of each object in making an overall size judgment. The judge was given the option to further divide a column if there appeared to be significant size differences among the objects in the column. The procedure was similar in format to those used for the shape and conceptual tasks. Judges generally took about 35-40 minutes to complete each domain, such that their entire session took about 90 minutes. For the detailed instructions, see Appendix E.

2.5.4. Classification and Rating Results

The object classification and rating results of Phase 2 are necessary for analyzing the FE performance of Phase 1 of this study. Hence, the results of Phase 2 will be presented first. The Phase 2 results include shape, semantic, and size classification, and shape similarity rating, and agreement.

2.5.4.1. Shape Similarity Results. Table 1 shows the Judge 1-Judge 2 (J1-J2) and the judge-average (Js-Avg) rating correlations, indicating a good level of agreement. The average (Avg) ratings were entered into the subsequent analyses reported below.

2.5.4.2. Shape and Semantic Classification Agreement Results for Master Sets. Table 2 displays the levels of agreement between judges (Js-Js), between judges and the composite (Js-C), and between semantic and shape composites (C-C), for each domain's master sets. (See section 2.5.3.5 for a description of the composite). There were multiple judges. Thus, rather than report all agreement scores, Js-Js and Js-C agreement scores are means based on all J-C pairings and non-redundant J-J pairings. As the Cramér's *V* scores indicate, J-J agreements are reasonably high, though not as high as the J-C agreements. This indicates that the composite in each instance is a representative "average" of the different judges' classifications.

Table 1

Interjudge Agreement for Shape Similarity Ratings (FE Output)

Domain	*df*	Mean[a]	J1-J2	Js-Avg
Long	409	3.46	.85	.96
Flat	390	3.36	.83	.96
Round	372	3.61	.88	.97
Straight	393	3.34	.87	.97

Notes. J1-J2 = correlation between Judge 1's and Judge 2's ratings on successive pairings of objects in FE output sequences. Js-Avg = the mean correlation between judges' ratings and the average. The *dfs* are based on the total number of item pairings in each domain.
[a]The mean of the average (Avg) shape similarity rating within the domain (range 1-7, where 1 = very low, and 7 = very high).

The shape and semantic classification results in Table 2 are for the specific level of categorization, which is shown in the object lists in the final Appendix, N. The general (or broader) level of classification composite is also presented prior to the object lists in Appendix N. There was near-perfect agreement between the specific and the general classification composite for shape (mean V = .995) and for semantic (mean V = .992). The mean agreement between specific shape and semantic classification composites, listed in Table 2, above, was V = .55, and between general semantic and general shape classification composites was V = .48. The general classifications were taken into consideration for the classification of the actual FE output.

The comparisons between the semantic and shape composites show that (a) this type of agreement is, as one should expect, much lower than the levels of agreement within those respective classification schemes; and (b) there is considerable overlap between semantic and shape category membership, as expected due to the natural confounding of those factors in the object population sampled. All of the agreement results were significant according to the ordinary χ^2 measure, except for the Long Shape Composite – Semantic Composite comparison. The apparent lack of significance in that case was likely only an artifact associated with the large number of single-item shape categories for Long. Removing the single-item categories from the χ^2 analysis revealed significant results, consistent with the semantic-shape comparisons for the other domains. Moreover, the Z-score calculations (Brennan & Light, 1974), using Hubert's (1977) formula, showed that this comparison was significant with those numerous single-item categories included in the analysis, Z = 14.48.

Table 2

Shape and Semantic Classification Agreement for Master Sets

	Domain	*df*	χ^2	*V*
Shape: J-J	Long	2923	7450	.78
	Flat	1200	4610	.81
	Round	1266	5892	.78
	Straight	1543	5771	.81
Shape: J-C	Long	4021	10826	.89
	Flat	1631	5804	.88
	Round	1998	7915	.87
	Straight	1699	7000	.86
Semantic: J-J	Long	1672	7331	.88
	Flat	1292	5640	.88
	Round	1936	8394	.81
	Straight	1554	5361	.77
Semantic: J-C	Long	1558	7689	.90
	Flat	1044	5688	.95
	Round	1936	9374	.86
	Straight	1620	6662	.84
Shape-Semantic: C-C	Long	2812	2771	.54
	Flat	1421	2170	.59
	Round	2464	3534	.53
	Straight	1763	2865	.54

Notes. J = judge, C = composite. For multiple judge-judge (J-J) or judge-composite (J-C) pairings, the *means* of the *df*, χ^2 and Cramér's *V* scores are reported. Means are based on all non-redundant J-J or J-C pairings for the relevant comparisons. Mean *df*s and χ^2 scores are rounded off to the nearest whole number.

2.5.4.3. Size Agreement Results. Table 3 shows that high levels of agreement were obtained between size judges (mean r = .96). There was also clearly significant association, as measured by the η statistic,[12] between shape and size, and between semantic category and size. This confirms the common sense expectation that there is substantial natural confounding between these different dimensions in the object populations sampled. For example, food items tend to be small, furniture items tend to be near the middle, and vehicles and houses tend to be at the upper end of the size scale. Some of the shape-size association is probably due to the semantic variable, and due to some size-shape similarities across semantic categories (e.g., balls and fruits/vegetables).

2.5.5. Scoring and Measurement of FE Output

2.5.5.1. Production. Item production was the number of items reported correctly within seven minutes. An item was scored as correct if it conformed to the general shape domain category. Note that 99.7% of the items were of correct shape. Where an item was reported a second time in the same list, the later instance (i.e., the item "repeat") was not counted as a correct listing. Repeats made up 1% of the total production. In the occasional instances where successive items in a run were extremely similar in both shape *and* semantic category (e.g., "…quarter, nickel, dime, penny…"), only two items in the run were counted toward the production total. This conservative adjustment was made in order to prevent the inflation of production and shape and semantic clustering scores, respectively. Only three kinds of these redundant runs occurred, namely, balls used in sports/games, coins, and paper money. Although 9 of the 22 participants had at least one such redundant run, the adjustment caused a reduction of only 2% of the total items produced in the experiment.

Table 3

Agreement for Size, Size-Shape, and Size-Semantic Classification: Master Sets

Comparison	Domain	*df*	*r*	η
Size: J-J	Long	250	.93	
	Flat	211	.97	
	Round	273	.96	
	Straight	241	.97	
Size: Js-Avg	Long	250	.98	
	Flat	211	.99	
	Round	273	.99	
	Straight	240	.99	
Size Avg - Shape C	Long			.72
	Flat			.71
	Round			.62
	Straight			.71
Size Avg – Semantic C	Long			.78
	Flat			.74
	Round			.63
	Straight			.82

Notes. J = judge, C = composite, Avg = Average, *r* = Pearson correlation, *df* = degrees of freedom for *r*; η = eta, a statistic for measuring nominal x interval association.[12]

2.5.5.2. Clustering. A *cluster* is here defined as a sequence of two or more consecutive items that are in the same category. Clusters consist of categorical repetitions. For example, consider the sequence *abccdefgggh*, where the different letters represent categories and the individual occurrences in the sequence represent items. In that example there are 11 items, 8 categories, and two clusters. There is one cluster for category *c*, and one cluster for category *g*. Category *c*, with two consecutive items, is repeated once, and thus contains one categorical repetition. Category *g*, with three consecutive items, is repeated twice, and thus contains two categorical repetitions. The total number of categorical repetitions in that sequence is 1 + 2 = 3.

The Adjusted Ratio of Clustering (ARC; Roenker, Thompson, & Brown, 1971) was used to measure clustering in FE output sequences (for formulae, see Appendix F). The ARC ranges from a maximum of +1 to a minimum of –1, and has a chance level of 0 (zero). Other researchers have previously used the ARC to measure clustering in FE data (e.g., Bond et al., 1985; Fiske, 1995; Plumert, 1994). Using the ARC requires only an output sequence containing distinct categories of items. It has been shown to be robust to unequal category sizes and unequal numbers of categories (Roenker et al., 1971), and is not artifactually correlated with list length (Murphy, 1979).

Shape category assignment was determined most strongly by the composite sort of the relevant master set and by the original FE participant's classification, but the shape similarity ratings for pairings of items, and the original FE participant's shape category and exemplar information, were also considered. In cases where the mean similarity rating for a pair of items was approximately 4 and there was not a clear categorical division between the items in the pair according to the classifications, at least one item that was ambiguous was treated as a filler item. This meant that the item could not score in a cluster, and was treated, for the purposes of ARC scoring, as a single item in an additional category. Problematic items such as repeats and errors were also treated the same way as filler items. Problematic items that occurred at the end of an output list were simply excluded.

Semantic category assignment was based on the composite semantic classifications of the master sets. Occasionally, some items could be members of more than one semantic category. For example, *poker chip* could be a *money* or a *game* item, but, for scoring, the item had to be assigned to only one of the categories. The FE participant's output list context was sometimes used to help resolve such ambiguities. For example, if the participant had listed game items in addition to poker chip, but not money items, then poker chip would be assigned as a game item. Items that could not be scored clearly as clustered or not clustered were treated as filler items in the scoring, as above.

In making use of the master list composite classification in scoring clusters versus switches versus fillers, both the general and specific levels of classification were taken into account (see Appendix N), in addition to the rating and classification information provided by the FE participant and the shape similarity judges. Because the sets of objects within an FE list are much more limited than the master list sets, the classifications in the latter tended to be more precise and specific than those in the former. Thus it was often the case that the general or broad level of classification was most relevant to the set of items a given FE list.

2.5.5.3. Agreement Between Shape Cluster Scoring and Shape Similarity Rating. The correlation between dichotomous shape (cluster = 1, switch = 0) and average shape similarity (range 1 to 7) for the relevant IRTs (N = 1536) was r(1534) = .85. This level of agreement also held for the subsamples (N = 1301, and N = 1387) that are described further in the IRT Results sections, below. For the N = 1301 subsample, the level agreement was r(1299) = .86. The correlation between the mean average shape similarity and mean shape dichotomous scores of the FE output lists (N = 44) was r(42) = .91. The transitions that were treated as fillers in the scoring of clustering, which occurred in the middle range on the average shape similarity scale, were scored as .5 on the shape cluster (1) versus switch (0) measure. When those ".5" transitions were included in the N = 1387 subsample analysis, making a three-level shape variable (0, .5, 1), the correlation with the average shape similarity scale measure remained, r(1385) = .85.

2.5.5.4. Size Difference. Using the average of the rankings provided by the size judges, a size score was assigned to each object in the FE lists for which a size score had been obtained. The size difference score for each inter-item transition in the FE output lists was determined simply by subtracting the size of object t_1 from the size of object t_2, and so on. Size difference was the absolute value of the difference between the two size scores. For example, if object t_1 was size 5 and object t_2 was size 6, the difference in size between those objects was an absolute value of 1. If both objects in the pairing had the same size score, then the size difference score would be zero (0). In the output sequences, cases were excluded where extremely large (e.g., planet) or extremely small items (e.g., axon) made a transition with an ordinary object, because their size relative to the size of objects on the ordinary everyday scale is not clear.

2.5.5.5. Interresponse Times (IRTs). The tape-recorded output retrieval sequences were recorded onto a computer hard-drive (sampling rate = 44100 Hz, sampling size = 16 bits) and compact disc for scoring. Wavesurfer software (Sjolander & Beskow, 2004) was used to provide a graphic visualization of the waveforms of the spoken words in the output sequences. An IRT was measured from the onset of one word utterance to the onset of the next. This was done by pointing and clicking on the beginning of one word, scrolling to the beginning of the next word in sequence, and reading off the interresponse time in milliseconds. For each word, care was taken to identify the onset of articulation as exactly as possible. In cases where the person voiced "a" or "an" or "uh" right at the onset of the word and joining the word utterance, the beginning of those utterances was treated as the beginning of the report for that item. In cases where there was a noticeable delay between the onset of one of those preceding utterances and the word onset, as indicated by a clear visible gap separating the respective waveforms of the "uh" (or "a," "an," etc.) and the word onset, the word onset was treated as the beginning of the report for that item.

In scoring IRTs—consistent with the scoring of item production and clustering—runs containing items that were highly redundant in both shape and semantic similarity were reduced to two items. In those cases, the average IRT within the run was used as the IRT. In addition, the average of shape similarity, average size difference scores, and common semantic category within the run were used in those cases. Again, this was to prevent the inflation of the magnitude of the effects of interest.

To examine the temporal effects associated with item clustering, IRTs were classified as shape cluster or switch and semantic cluster or switch. This classification scheme is based on that used by Schreuder, Flores d'Arcais, and Glazenborg (1984; also see Flores d'Arcais, Schreuder, & Glazenborg, 1985) in their studies of object shape similarity effects in the processing of object nouns in lexical decision and naming tasks. The same scoring of FE output item pairings for the shape and semantic clustering (ARC) measurement was used for assigning items to this four-class scheme. IRTs were classified as follows: Semantic Switch-and-Shape Switch (SemSw—ShapeSw, e.g., *ring, pole*); Semantic Switch-and-Shape Cluster (SemSw—ShapeCl, e.g., *ball, grapefruit*); Semantic Cluster-and-Shape Switch (SemCl—ShapeSw, e.g., *cup, plate*); Semantic Cluster-and-Shape Cluster (SemCl—ShapeCl, e.g., *quarter, penny*).

2.6. FE Results

2.6.1. Production

As shown in Table 4, participants produced a mean of 36.7 items (the range of

Table 4

Mean Numbers of Items Produced in FE

Trial		Long	Flat	Round	Straight	*M*
1st	*M*	**40.5**	**39.9**	**30.5**	**34.8**	**36.3**
	n	4	7	6	5	
	SE	1.66	4.98	2.00	3.12	
2nd	*M*	**37.4**	**30.3**	**40.4**	**38.3**	**37.6**
	n	7	4	5	6	
	SE	5.39	2.10	4.43	4.74	
	M	**38.5**	**36.4**	**35.0**	**36.7**	**36.7**

persons' mean production was 22.5 to 56). Production did not increase significantly on the second as compared to the first trial, $t(21) = -0.3$, $SE = 2.43$, $p = .75$ (two-tailed), nor did it differ significantly between domains or domain pairs (the range of two-tailed *p*-values was .42 to .94).

2.6.2. Shape and Semantic Clustering

Table 5 shows the significant clustering by shape ($M = .29$) and semantic ($M = .42$) categories. There was no significant change from first to second trial for either of these types of clustering [both $ps > .58$ (two-tailed)]. There was greater mean semantic clustering in Long than in Flat, [$t(10) = 2.74$, $SD = 0.173$, $p = .021$ (two-tailed)], but otherwise there were no significant differences among the domains for mean shape or mean semantic clustering, respectively. Overall, semantic was higher than shape clustering, as shown on Trial 1, $t(21) = 3.52$, $SE = 0.038$, $p = .002$ (two-tailed), and on Trial 2, $t(21) = 4.88$, $SE = 0.028$, $p = .00008$ (two-tailed).

As indicated by the *t*-values displayed in Table 5, mean shape and semantic clustering were significantly above chance in each of the eight conditions (the one-tailed *p*-values ranged from .032 to .000013). Out of the 44 individual shape and semantic ARC scores, respectively, none (0) was below chance.

2.6.3. Clustering Density

Clustering density was measured as the number of cluster transitions divided by the total number of transitions. To examine the density of clustering over stages of output, the pooled data ($N = 1301$) were divided into approximately equal successive normalized fifths of output. (For a description of the normalizing procedure, see section 2.6.4.1, after Table 6). On the first trial, the density of shape and semantic clustering in

Table 5

FE Shape and Semantic Clustering (ARC)

Trial		Long	Flat	Round	Straight	*M*
			Shape ARC			
1st	*M*	**.29**	**.30**	**.32**	**.20**	**.28**
	n	4	7	6	5	
	t	10.29	8.44	4.46	3.05	
	SE	0.028	0.035	0.071	0.066	
2nd	*M*	**.32**	**.20**[a]	**.38**	**.26**	**.30**
	n	7	4	5	6	
	t	7.12	2.81	5.78	10.34	
	SE	0.046	0.070	0.066	0.026	
	M	**.31**	**.26**	**.34**	**.24**	**.29**
			Semantic ARC			
1st	*M*	**.50**	**.38**	**.39**	**.43**	**.41**
	n	4	7	6	5	
	t	12.29	11.43	5.83	7.01	
	SE	0.041	0.033	0.067	0.061	
2nd	*M*	**.48**	**.28**[a]	**.50**	**.42**	**.43**
	n	7	4	5	6	
	t	8.79	3.03	6.23	5.67	
	SE	0.054	0.091	0.081	0.075	
	M	**.48**	**.34**	**.44**	**.42**	**.42**

Notes. Number of participants per cell is represented by *n*. Maximum ARC = 1; chance ARC = 0. The *t*-scores indicate the extent to which the ARC scores are above chance (0).
[a]These two low means are each due to one low-outlying score in the respective analyses.

the first fifth of output was relatively high, then dropped for the second and third fifths, then rose slightly across the fourth and final fifths of output. On the second trial, cluster density was relatively uniform across fifths of output, but rose slightly in the fourth fifth.

2.6.4. IRT Results

The distribution of IRTs was highly positively skewed, as shown in Figure 1. There were 1536 IRTs, excluding those that bordered a repeated item (n = 35) and other problematic cases (n = 8), such as those transitions involving an item that could not be classified by shape, or which involved an unknown or unclear specific shape based on the information given by the FE participant. Of the 1536 IRTs, 1387 (90%) were used in many of the analyses below, which include only IRTs below 25 s, and which make use of the full shape similarity scale. In the N = 1387 sample, the slowest IRT was 24.940 s, and the fastest was 0.476 s. The mean and median IRTs were 8.04 s (SD = 6.277) and 6.48 s, respectively. In analyses below which are restricted to the dichotomous shape cluster-versus switch transitions (as determined in scoring the ARC), and which also include only IRTs below 25 s, the N = 1301, which constitutes 85% of the 1536 IRTs.

There are justifications for excluding the longer-than-25 s IRTs. One is that the occurrence of the long IRTs suggests that participants may have, temporarily, drifted off task. Another is that item content appears to be uncorrelated with IRT for durations 25 s and above. For example, semantic transition (switch = 0, cluster = 1) is unrelated to IRT duration in the longer-than-25 s subsample [r(147) = -.001]. That is unusual, given that the association between the semantic variable and IRT durations is well-established by previous research (as was discussed previously), though it may not be unusual to find no relation at longer-than-25 s durations. There was also no relation between IRT duration

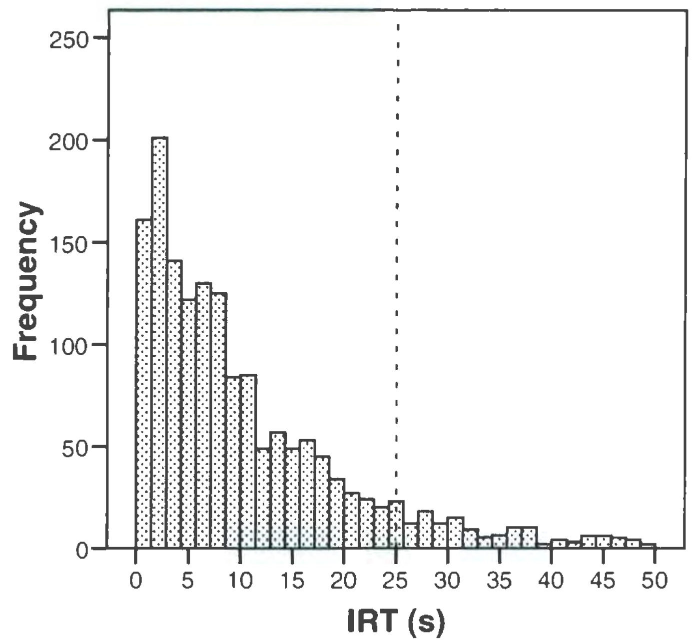

Figure 1. Distribution of IRT durations (s), with 25 s cut-off mark. [The tail of IRTs that were 50 s or longer (1% of the IRTs) has been omitted from the graph for presentation].

and (a) shape similarity [$r(147) = .017$] or (b) size difference [$r(138) = -.050$] in the longer-than-25 s subsample. (These non-relationships with IRT for shape, size, and semantic variables were also found when the sample was restricted to the 25 s to 35 s duration range). The problem is that, at time durations longer than approximately 25 s, tests pertaining to pairwise inter-item content may not be possible. Item content in FE may genuinely have no detectable effect at such long time durations, which is suggested by the above non-relationships. (Nonetheless, the statistically significant results reported below remain when the longer-than-25 s IRTs are included in the analyses).

Two different remedial actions were taken in order to deal with the positive skew. First, for most of the analyses reported below, unless otherwise noted, only IRTs that were shorter than 25 s were included. Second, for correlational and regression analyses (reported in later sections, below), IRTs were transformed successfully by using $\log_e(\text{IRT}_i + 0.525)$, where $\log_e$ is the base e logarithm, represented (below) by the abbreviation Ln. (The constant 0.525 was used simply to maintain positive numbers).

2.6.4.1. Shape and Semantic Cluster Versus Switch IRTs. To examine the shape clustering versus switch effect on IRT durations, IRTs were divided into the four-class scheme based on Schreuder et al. (1984), as described above (at the end of section 2.5.5.5). The present analysis included only IRTs that were (a) under 25 s, (b) classified as shape cluster or switch (not as fillers) in the ARC scoring, and (c) not repeats or other problematic cases. With these restrictions, the remaining *N* was 1301. The frequencies (and percentages) of IRTs in each of the four classes were as follows: SemSw—ShapeSw = 533 (41%), SemSw—ShapeCl = 199 (15.3%), SemCl—ShapeSw = 241 (18.5%), SemCl—ShapeCl = 328 (25.2%). In preparation for assessing the shape and semantic clustering versus switching effects, each participant's mean IRTs for each of the four classes of inter-item transitions were obtained from IRTs pooled from both of his or her FE domain outputs (i.e., Trial 1 and Trial 2) to obtain adequate numbers of IRTs per cell in as many cells as possible. The mean scores are summarized in Table 6.

Participants' (*N* = 22) mean IRTs of the four types were then entered into a two-factor within-subjects ANOVA, with categories (shape or semantic) as the factors and inter-item transition type (cluster or switch) as the levels of each factor. (There was one empty cell out of four cells for the 10^{th} participant. It was replaced with the mean for that cell across participants for the present analysis. Another analysis, which excluded the 10^{th} participant, also found the same pattern of significant results for the effects of interest here). Additional analyses on participants' median IRTs also showed the significant effects, though these effects were slightly less than those for the mean IRT data.

There was a significant main effect for shape. Participants' mean shape cluster IRTs were significantly shorter than their mean shape switch IRTs, $F(1, 21) = 32.89$, $MSE = 2.144$, $p = .00001$, $\eta^2 = .61$. The main effect for semantic transition (cluster vs. switch) was also significant, $F(1, 21) = 181.59$, $MSE = 2.246$, $p = 8.34 \times 10^{-12}$, $\eta^2 = .90$. The interaction was not significant, $F(1, 21) = 0.001$, $MSE = 3.294$, $p = .98$.

Specific comparisons were made to examine the predicted shape effect at semantic switch and cluster transitions, respectively. At semantic *switch* transitions (SemSw), shape cluster were significantly shorter than shape switch mean IRTs, $t(21) = 3.18$, $SE = 0.566$, $p = .002$ (one-tailed), $\eta^2 = .33$. At semantic *cluster* transitions (SemCl), shape cluster were significantly shorter than shape switch mean IRTs, $t(21) = 4.26$, $SE = 0.418$, $p = .0002$ (one-tailed), $\eta^2 = .46$. These results show that shape clustering is associated with reductions in mean IRT durations whether or not it occurs with semantic clustering.

Next, the semantic effect was examined at shape switch and cluster transitions, respectively. At shape *switch* transitions (ShapeSw), semantic cluster were significantly shorter than semantic switch mean IRTs, $t(21) = 8.91$, $SE = 0.484$, $p = 7.03 \times 10^{-9}$ (one-tailed), $\eta^2 = .79$. At shape *cluster* transitions (ShapeCl), semantic cluster were significantly shorter than semantic switch mean IRTs, $t(21) = 8.28$, $SE = 0.519$, $p = 2.38 \times 10^{-8}$ (one-tailed), $\eta^2 = .77$. Thus, semantic clustering is associated with reductions in mean IRT durations whether or not it occurs with shape clustering.

In addition, at shape *switch* transitions, semantic cluster mean IRTs were significantly shorter than shape cluster mean IRTs at semantic *switch* transitions, as shown in the SemCl—ShapeSw versus SemSw—ShapeCl comparison, $t(21) = 4.98$, $SE = 0.505$, $p = .00006$ (two-tailed), $\eta^2 = .54$. This confirms that the semantic cluster speed advantage was larger than the shape cluster speed advantage.

The advantage of clustering over switching was largest in the comparison between double cluster and double switch transitions. The comparison was highly significant between SemCl—ShapeCl and SemSw—ShapeSw, $t(21) = 16.07$, $SE = 0.379$, $p = 1.42 \times 10^{-13}$ (one-tailed), $\eta^2 = .93$.

It is well-established that IRTs tend to increase in duration with increasing output position in people's retrieval sequences (Wixted & Rohrer, 1994). That trend was also

Table 6

Participants' Mean Shape and Semantic Cluster and Switch IRTs (N = 22)

Category and Transition		*M*[a]	*SE*	*Mdn*[b]	*SE*
Shape	Switch	8.66	0.258	7.44	0.324
	Cluster	6.87	0.331	5.79	0.354
Semantic	Switch	9.92	0.289	9.15	0.352
	Cluster	5.61	0.308	4.07	0.376
SemSw—ShapeSw		10.82	0.322	10.00	0.441
SemSw—ShapeCl		9.02	0.472	8.30	0.589
SemCl—ShapeSw		6.50	0.382	4.87	0.430
SemCl—ShapeCl		4.72	0.362	3.27	0.398

Notes. [a]Means of participants' mean IRTs (s). [b]Means of participants' median IRTs (s).

shown in the present data. To examine the possible influence of output position on the above results, the four-class analysis (see above) was conducted on the normalized rank output positions of the IRTs. The normalized rank output position for each IRT was calculated by the formula ROP / n + 1, where ROP is the rank output position of the IRT in the person's recall output for a given trial, and n is the number of IRTs in the trial. The persons' four types of mean and median normalized ROPs were then obtained in the same way as their mean and median IRTs were obtained. There were no main effects for shape or semantic transition, no interactions, and no significant specific pairwise differences for either mean or median normalized ROPs. (The results remained non-significant whether the 10^{th} participant's missing score was replaced with the appropriate cell mean or excluded from the analyses). Therefore, the shape cluster speed advantage, revealed in the four-class analysis of IRTs above, cannot be attributed to earlier output position.

There were not sufficient IRTs for assessing the possible differences between domains and trials, respectively. However, graphical inspection of the means for those conditions does suggest the presence of some differences. The pattern of a shape cluster reduction in mean IRT durations across semantic switch and cluster conditions occurred in each FE domain (see Figure 2). This shape cluster advantage appears stronger in Long and Round domains, but weaker at semantic switch transitions in the Flat and Straight domains. The stronger shape cluster IRT effect for Long and Round is consistent with the higher mean shape clustering scores for those domains relative to Flat and Straight.

There also appears to be a larger shape cluster effect on IRTs in the first as compared to the second trial domain (see Figure 3). However, the magnitude of the

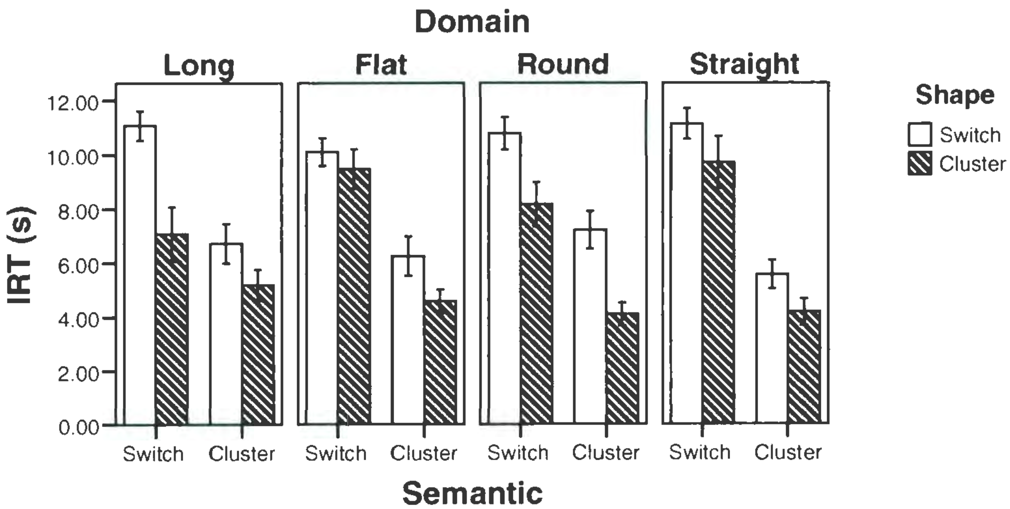

Figure 2. Mean IRTs for shape and semantic cluster versus switch transitions for each FE domain (N = 1301). Error bars ± 1 *SE*.

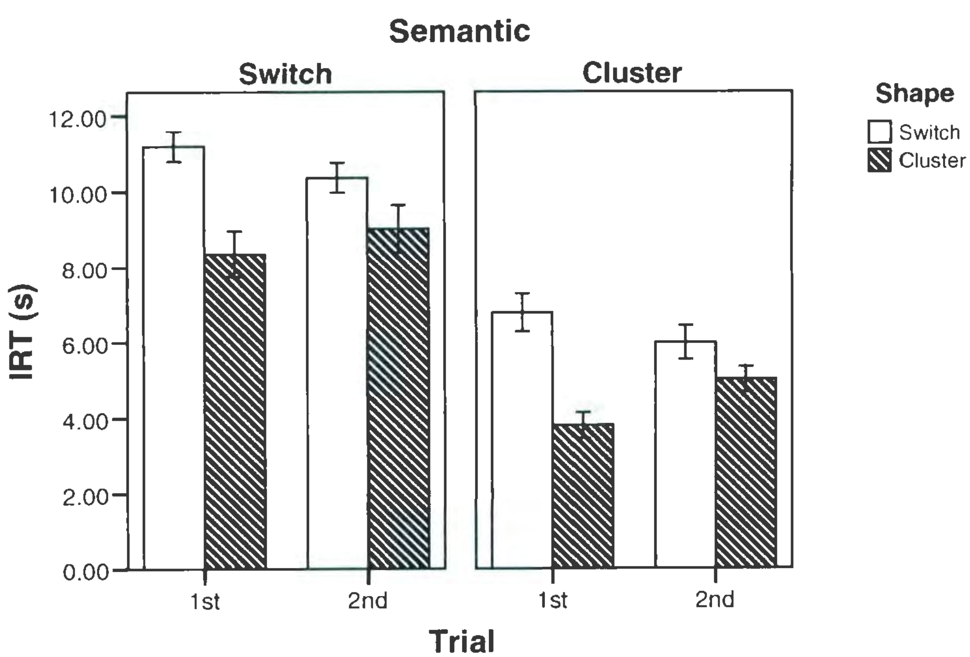

Figure 3. Mean IRTs for shape and semantic cluster versus switch transitions, at Trial 1 versus Trial 2 (N = 1301). Error bars ± 1 *SE*.

semantic cluster effect on IRTs (not displayed directly here) was consistent for the first and second domain trials.

There were not sufficient IRTs per participant to incorporate the inter-object size difference in the above statistical analyses. Nonetheless, the shape cluster speed advantage observed at semantic switch transitions is not attributable to a size difference effect, because there was no size difference effect in that condition. The relation between IRT duration and the variables of interest is dealt with in the next sections.

2.6.5. Relationships Between IRT Durations and the Variables of Interest

In this section, I will focus on the relationship between IRT durations and each of the variables of interest, including shape similarity, semantic transition (cluster versus switch), and size difference, but also output position and verbal response length. I will present evidence regarding the relationships between these variables, especially as they relate to shape similarity.

2.6.5.1. Shape Similarity Levels. The mean and median shape similarity scores for the $N = 1387$ sample were 3.52 ($SD = 1.920$) and 3.5, respectively. Table 7 shows the levels of shape similarity at shape cluster versus switch transitions for the $N = 1301$ sample. When listing object names, FE participants tended to show shorter IRT durations between objects that were more highly similar in shape, and longer IRT durations between objects that were less similar in shape. This trend is visible in Figure 4, which displays the median IRTs for each level of average shape similarity. Table 8 lists the mean and median IRT scores for each level of shape similarity.

Table 7

Shape Similarity Levels at Shape Cluster and Switch Transitions (N = 1301)

Transition	*n*	*M*	*SD*	*Mdn*
Switch	774	2.14	1.030	2.0
Cluster	527	5.56	0.978	5.5

Note. On the shape similarity scale, 1 = very low, and 7 = very high.

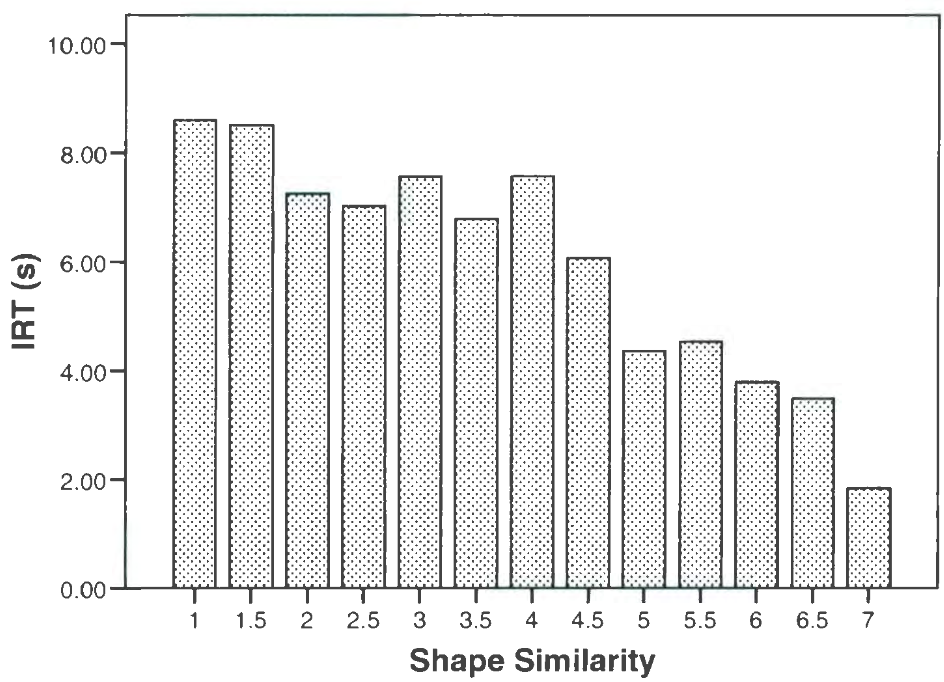

Figure 4. Median IRTs (s) at each level of shape similarity (N = 1387). On the shape similarity scale, 1 = very low, and 7 = very high.

Table 8

IRTs (s) For Each Level of Shape Similarity (N = 1387)

Shape Similarity	*n*	*M*	*SD*	*Mdn*
1.0	189	10.24	6.299	8.60
1.5	163	9.91	6.586	8.51
2.0	128	9.10	6.547	7.25
2.5	118	8.42	6.025	7.02
3.0	78	8.70	6.180	7.56
3.5	90	9.05	6.799	6.79
4.0	84	8.69	6.215	7.57
4.5	106	7.99	6.218	6.07
5.0	115	6.54	5.588	4.36
5.5	81	5.69	4.804	4.53
6.0	74	6.00	5.719	3.80
6.5	90	5.41	5.318	3.49
7.0	71	3.25	3.218	1.85

Note. On the shape similarity scale, 1.0 = very low, and 7.0 = very high.

2.6.5.2. Semantic Category. As reported above, and consistent with previous research that used other types of FE task domains, semantic cluster IRTs tended to be much shorter than semantic switch IRTs in these FE tasks. For semantic cluster versus switch transitions, Tables 9, 10, and 11 show the descriptive statistics for IRTs, shape similarity levels, and size difference levels, respectively. The mean semantic transition score (where cluster = 1 and switch = 0) for the $N = 1387$ sample was 0.43 ($SD = 0.496$).

Table 9

IRTs (s) at Semantic Cluster and Switch Transitions (N = 1387)

Transition		*n*	*M*	*SD*	*Mdn*
	Switch	785	10.19	6.401	9.05
	Cluster	602	5.24	4.850	3.60

Table 10

Shape Similarity Levels at Semantic Cluster and Switch Transitions (N = 1387)

Transition		*n*	*M*	*SD*	*Mdn*
	Switch	785	2.95	1.717	2.50
	Cluster	602	4.28	1.909	4.50

Note. On the shape similarity scale, 1 = very low, and 7 = very high.

Table 11

Size Difference Levels at Semantic Cluster and Switch Transitions (N = 1358)

Transition		*n*	*M*	*SD*	*Mdn*
	Switch	760	2.28	1.800	2.00
	Cluster	598	1.02	1.089	0.75

Note. Zero (0) = no size difference and higher numbers indicate higher levels of size difference.

2.6.5.3. Size Difference. IRT durations tend to increase as the inter-object size difference increases, as shown in Figure 5. For presentation purposes, levels of size difference are binned into increments of 1. The original size difference increments were generally of 0.25 throughout most of the distribution, but above the level of 7, there were fewer scores and some coarser increments of .5. There were 33 increments of size difference, including zero (0). The binning scheme was as follows: 0 (no binning; includes scores of zero); 1 included .20 to 1; 2 included 1.25 to 2; and so on. Table 12 shows the mean and median IRT for each level of binned size difference. There were more inter-item transitions with low as compared to higher size difference.

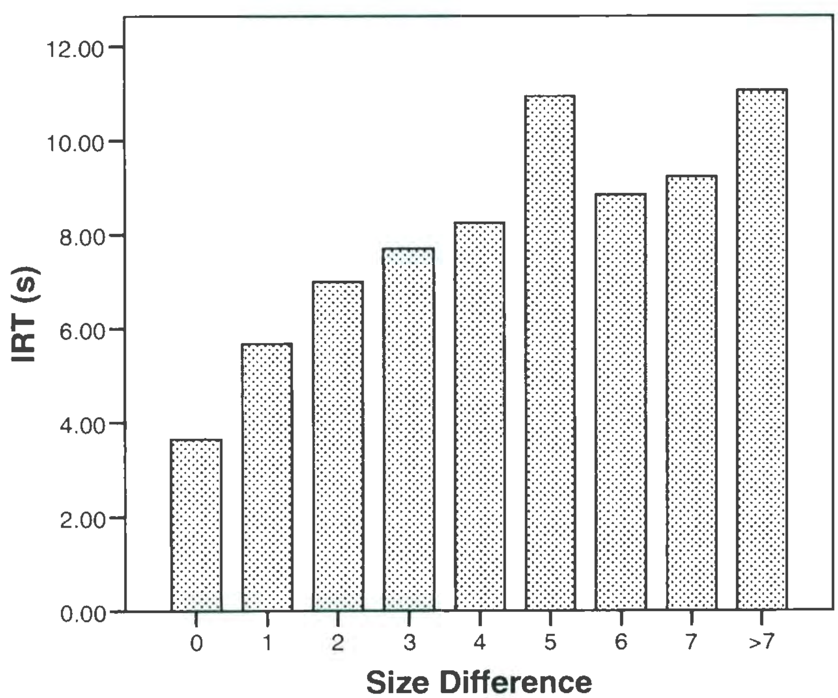

Figure 5. Median IRTs (s) and binned inter-object size difference (N = 1358).

Table 12

IRTs (s) for Levels of Inter-Item Size Difference (N = 1358)

Size Difference	*n*	*M*	*SD*	*Mdn*
0	200	5.65	5.365	3.64
1	460	7.15	5.995	5.68
2	285	8.61	6.368	7.00
3	170	8.73	5.948	7.70
4	107	10.34	6.817	8.24
5	70	11.23	6.721	10.94
6	34	9.64	6.797	8.84
7	22	11.20	6.561	9.22
>7	10	11.12	5.730	11.07

Note. A size difference of 0 indicates no size difference, and a size difference of >7 indicates very high size difference. Scores are binned into size difference increments of 1, except for the value of zero (0), which includes only scores of zero (i.e., no size difference).

The relation between IRT duration and size difference occurs across levels of shape similarity (see Figure 6). Likewise, the relation between IRT duration and shape similarity occurs across levels of size difference (see Figure 7). For other analyses, size difference transitions were divided by a median split into those of low size difference (below the size difference median) and those of high size difference (above the size difference median). The low size difference (below median) mean size difference was 0.5 (SD = 0.42), and the high size difference (above median) mean size difference was 3.0 (SD = 1.48) for the N = 1273 subsample (see Figure 8). For the N = 1358 sample, the mean and median size difference scores were 1.73 (SD = 1.650) and 1.25.

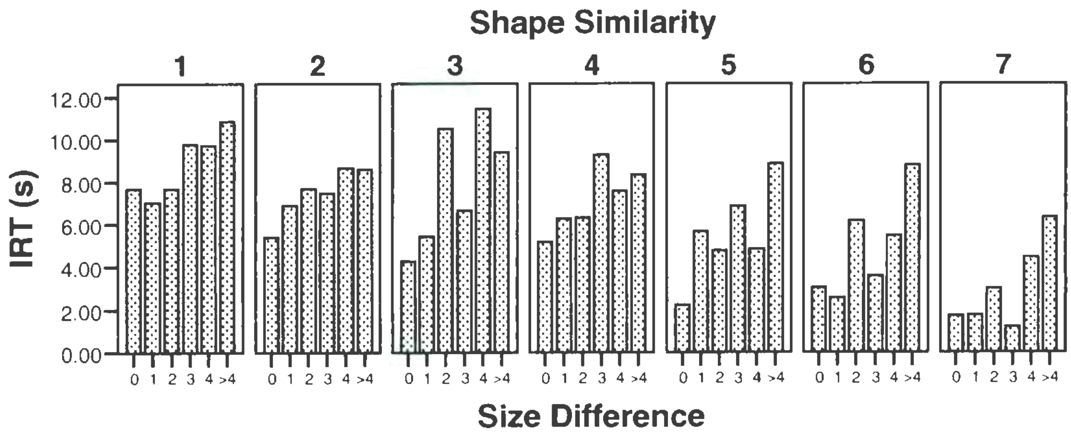

Figure 6. Median IRT (s) and size difference (bins 0 to >4) across levels of shape similarity (binned: 1 to 7, where 1 includes 1 and 1.5; 2 includes 2 and 2.5, etc., and 7 includes only 7) (N = 1358).

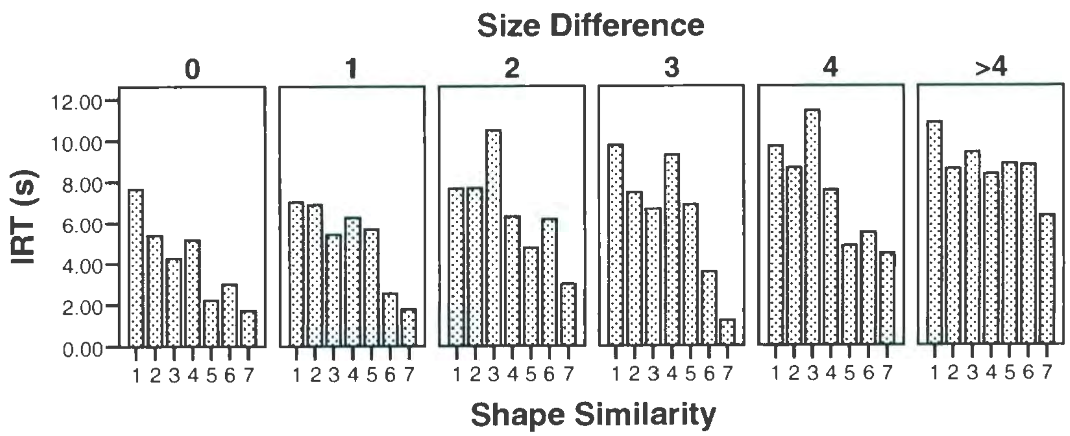

Figure 7. Median IRT (s) and shape similarity (7 bins) across levels of size difference (bins 0 to >4) (N = 1358).

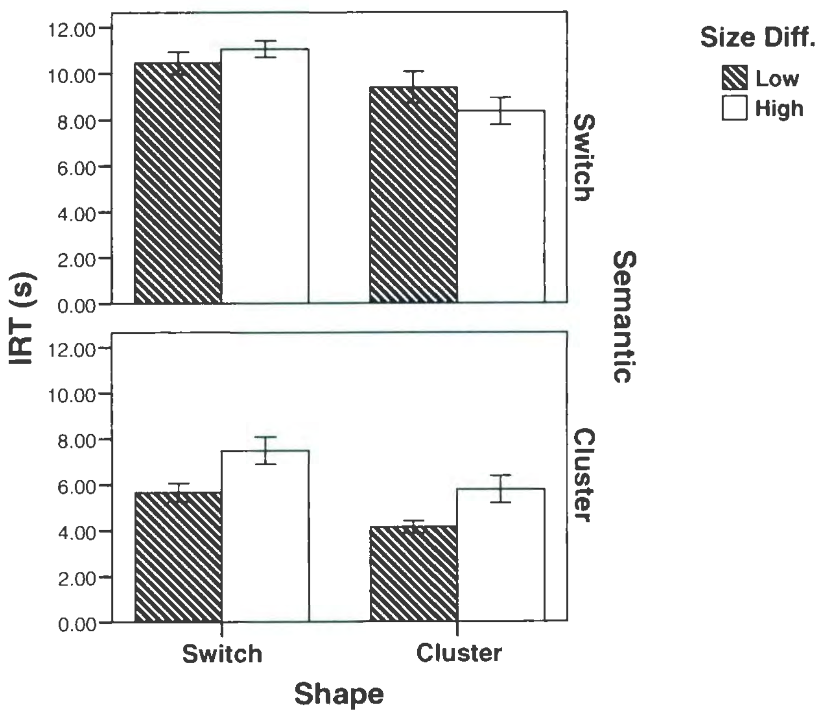

Figure 8. Low versus high size difference IRTs (s) at shape and semantic cluster versus switch transitions (N = 1273). Error bars ± 1 *SE*. For this sample, the mean size difference for the low size difference transitions is 0.5 and for the high size difference transitions is 3.0.

The Figure 8 shows that the speed advantage for transitions of low size difference was limited to semantic cluster transitions. Figure 9 shows that there are more transitions of higher shape similarity within the semantic cluster subsample, and more transitions of lower shape similarity within the semantic switch subsample. There are more transitions of low size difference for the semantic cluster subsample and more transitions of high size difference for semantic switch subsample. In addition, there are relatively more transitions of low compared to high size difference as shape similarity increases within the semantic cluster subsample.

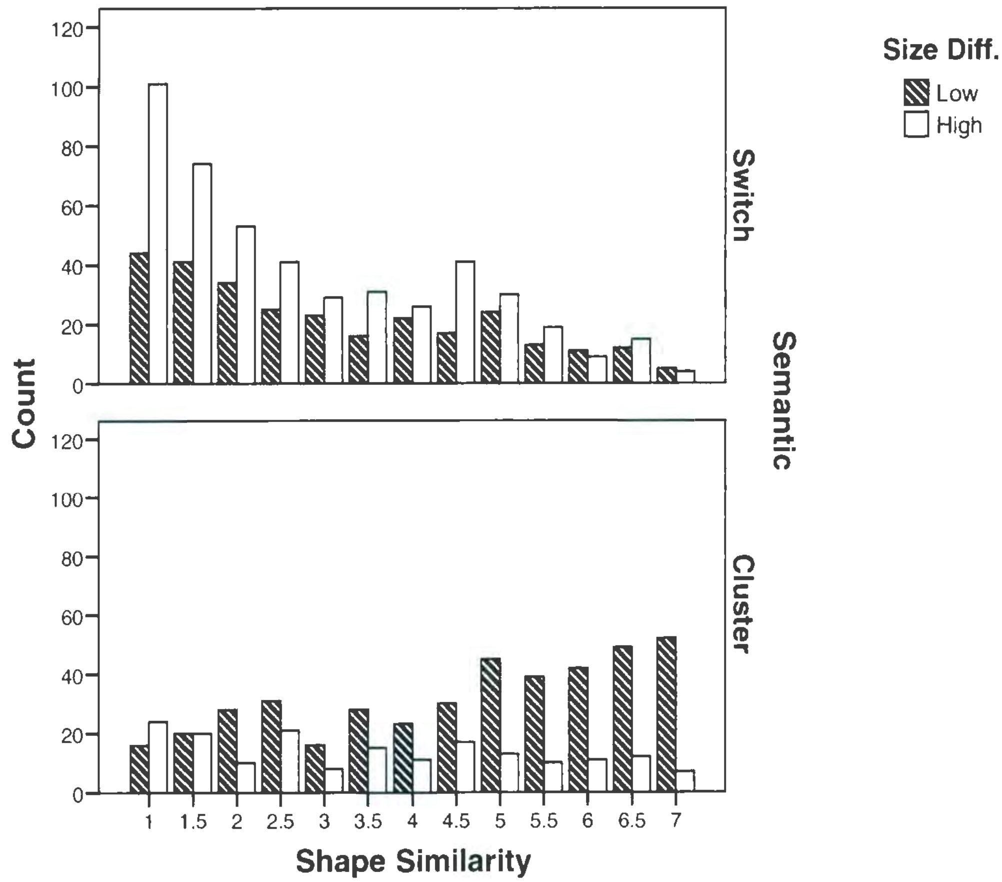

Figure 9. Frequencies of low versus high size difference transitions at each level of shape similarity and at semantic cluster versus switch ($N = 1358$). For this sample, the mean size difference for the low size difference transitions is 0.5 and for the high size difference transitions is 3.0. For shape similarity, 1 = very low and 7 = very high.

2.6.5.4. IRT Output Position. IRT durations increased most dramatically over the first approximately 10 IRT output positions, continued to increase somewhat until approximately the 20th to 25th output positions, and then increased little after that (see Table 13). Table 13 shows the IRTs for the first 10 output positions, followed by bins of 5 output positions, ending with binned positions 46+. For the $N = 1387$ sample, the mean and median output positions were 19.27 ($SD = 12.957$) and 17.0, respectively.

Table 13

IRT Durations Across IRT Output Positions (N = 1387)

Output Position	*n*	Cum. %	*M*	*SD*	*Mdn*
1	44	3.2	1.45	0.943	1.15
2	44	6.3	2.53	2.464	1.55
3	41	9.3	2.95	2.608	1.91
4	42	12.3	4.00	4.216	2.62
5	42	15.4	4.76	4.309	3.53
6	43	18.5	5.36	5.215	3.02
7	41	21.4	7.06	5.087	6.47
8	43	24.5	6.38	5.288	5.22
9	43	27.6	6.31	3.990	6.10
10	43	30.7	8.47	6.259	7.83
Bin					
15	200	45.1	8.63	5.825	7.33
20	186	58.5	9.47	6.757	7.46
25	162	70.2	9.67	6.114	8.58
30	143	80.5	10.14	6.575	8.85
35	93	87.2	10.09	7.000	7.94
40	74	92.6	9.12	6.877	7.02
45	48	96.0	8.87	5.944	7.94
46+	55	100.0	9.43	5.681	8.30

Notes. Cum. % = cumulative percent. IRTs are binned in output position ranges as follows: 15 includes scores from output positions 11 to 15; 20 includes scores from output positions 16 to 20; and so on; while the last bin, 46+, includes positions 46 to 61.

The conspicuous reduction in IRT durations that begins at about position 40 and continues to at least position 45 (see Table 13) is attributable to the fact that those participants who produced more responses than the mean of 36.7 items had shorter IRTs overall and contributed the higher output positions, whereas those who produced less IRTs necessarily had longer IRTs and did not contribute the higher output positions. To adjust for these differences among participants, output positions were converted to normalized ranks using the formula ROP / n + 1, as described above. Thus, an output position of 21 in participant A's output sequence of n = 21 IRTs has a normalized rank of .9545, while an output position of 21 in participant B's output sequence of n = 42 IRTs has a normalized rank of .4883. Normalized ROP was then transformed by adding the constant 1 and using the Ln transformation in order to increase, slightly, the amount of explained variance in IRT for subsequent correlation and regression analyses.

The speed advantages for shape (Figure 10) and semantic clustering (Figure 11), and for low versus high size difference (Figure 12), occur across all stages of output.

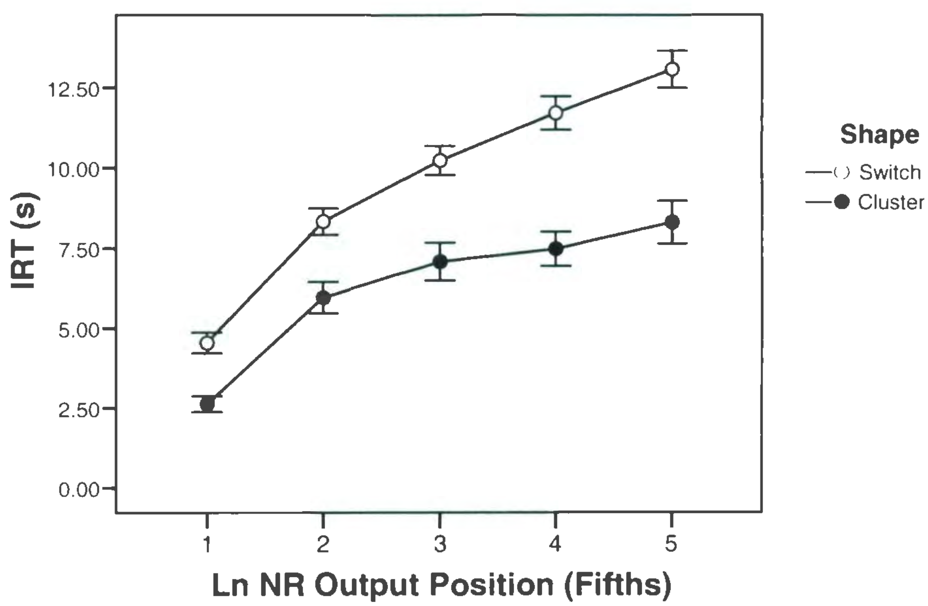

Figure 10. Shape cluster versus switch mean IRTs across successive fifths of Ln normalized rank output position (N = 1301). Error bars ± 1 SE.

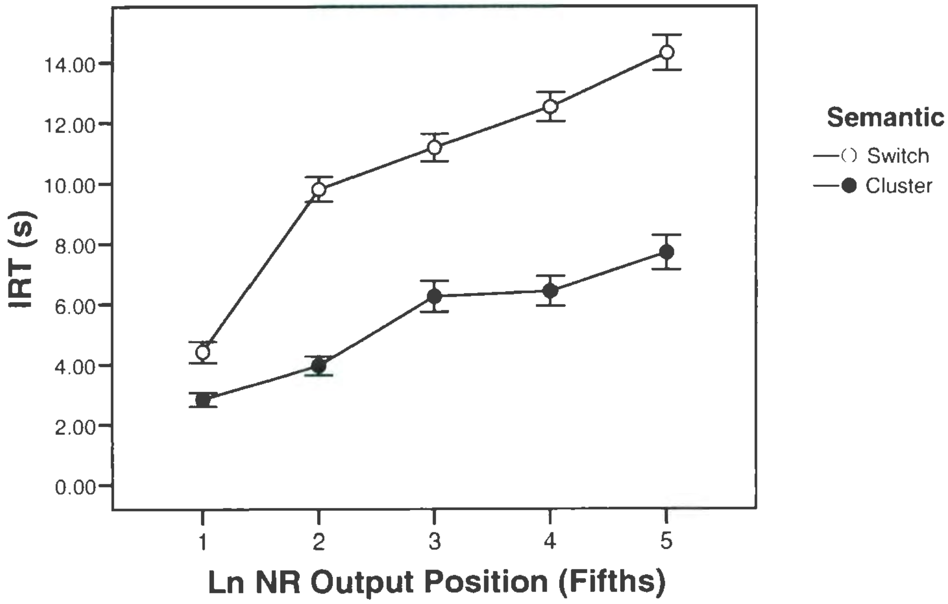

Figure 11. Semantic cluster versus switch IRTs across successive fifths of Ln normalized rank output positions (*N* = 1301). Error bars ± 1 *SE*.

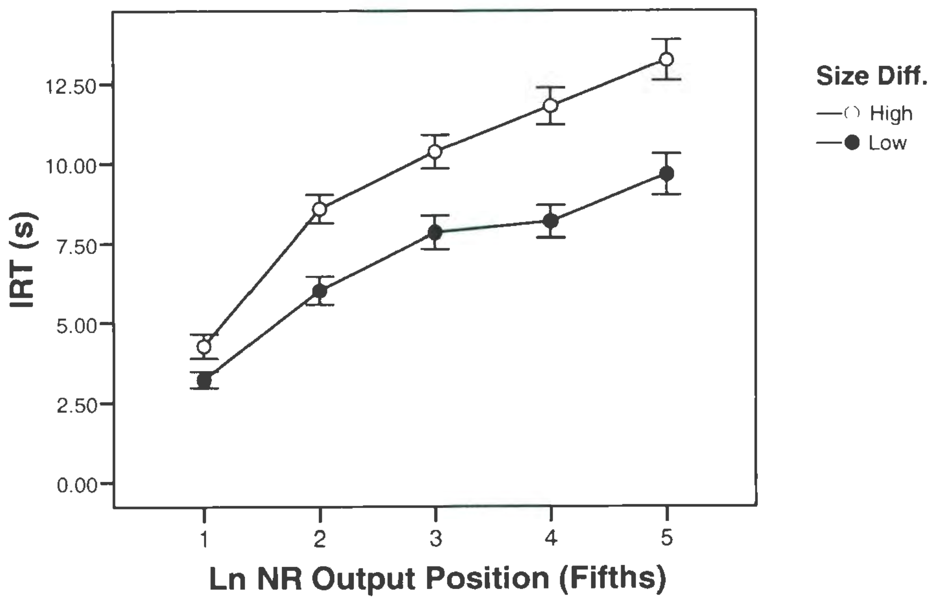

Figure 12. Low versus high size difference IRTs across successive fifths of Ln normalized rank output positions (*N* = 1358). Error bars ± 1 *SE*.

For the data presented in Figures 10-12, Ln normalized rank output positions were binned into successive, approximately equal fifths. Note that the effects for shape similarity, semantic transition, and size difference increase after the first phase of output.

2.6.5.5. Verbal Response Length. This variable was a composite of the number of syllables and the number of words in a response. Table 14 shows that as verbal response length increases, IRT durations also increase, but above a verbal response length of approximately 3, there is not a consistent increase in IRT durations. There were no significant correlations between verbal response length and shape similarity, size difference, or semantic transition (see next section). For the *N* = 1387 sample, the mean and median verbal response length scores were 2.07 (*SD* = 0.946) and 1.81, respectively.

Table 14

Verbal Response Length and IRT (N = 1387)

Verbal Response Length	*n*	*M*	*SD*	*Mdn*
1	139	5.95	5.40	4.30
1.5	399	7.41	6.47	5.75
2	315	8.01	6.29	6.63
2.5	239	9.16	6.69	7.51
3	133	9.01	6.14	7.79
4	120	8.73	5.45	7.52
5	23	10.51	5.31	10.07
6	10	7.14	2.87	7.93
>6	9	10.83	4.37	8.65

Note. Binning of verbal response lengths is as follows: 1 is an original increment that is not binned; 1.5 includes 1.25 and 1.5; 2 includes > 1.5 to 2; and so on; and >6 includes 6.25 to 9.

Verbal response length was added to the correlation and regression analyses (below) primarily to account for speech production and the cognitive processing costs associated with it. Words and phrases that are longer should generally be associated with longer IRTs. Verbal response length, as a composite variable, takes into account (a) the number of words in the response (usually one or two, sometimes more) and (b) the number of syllables in the response. These two variables were measured at two points: (1) in the IRT_i, and (2) in the IRT_{i+1}. The composite variable was simply the mean of (a) and (b) at levels (1) and (2); i.e., a mean of a1, a2, b1, b2. The composite was used in these analyses because it was more highly correlated overall with IRTs than were a1, a2, b1, and b2 individually or paired.

2.6.6. Correlation and Regression Analyses

Correlation and linear regression analyses were carried out to estimate the strength of the relation between inter-object shape similarity and IRT duration, while taking into account the role of the other variables. To prepare the data, transformations were carried out to normalize the distributions of some of the variables and to linearize their relation to IRT. Size difference, verbal response length, normalized rank output position, and IRT were each transformed using the base *e* log transformation (Ln). Before this transformation, a constant was added to variables that had some values less than 1 [IRT (+.525), size difference (+1), normalized rank output position (+1)] so that the scores would remain positive.

Autocorrelation in the residuals for IRT was another important concern. For most FE tasks, one should expect positive autocorrelation to be present in IRT sequences because of the significant clustering, which involves runs of IRT residuals that are of the same sign and of similar magnitude. Inspection of autocorrelation function and partial autocorrelation function plots of the pooled IRTs (N = 1387) indicated the presence of autocorrelation, beyond the 95% confidence limits, up to 3 to 4 lags. Multiple regression analyses on the pooled IRTs also showed that, while the first-order (lag-1) autocorrelation was reduced by the inclusion of the variables of interest into the regression model, it was still significant (Durbin-Watson obtained = 1.601).

To deal with the lag-1 and higher-order autocorrelation, adding more variables (besides those included below) to the regression equation was not feasible. Therefore other methods were considered. The two methods settled upon were, as suggested by Frei, "...to select only points that are far enough apart (in time) so that they are uncorrelated; or average values over a time period that is longer than the autocorrelation time scale" (Frei, 2003). Thus, one method, used below, makes use of mean (or median) scores for the variables for each output list, and the other includes in the analysis every *i*th case at a lag long enough that the autocorrelation is no longer significant.

2.6.6.1. Regression Analysis on Mean Scores. For the first correlation analysis, the mean score on each variable from each participant's FE output sequence (22 x 2 = 44) was used. The mean scores for the following variables were included: Ln IRT, shape similarity, shape ARC (clustering), semantic transition (0 = switch, 1 = cluster), semantic ARC, and Ln size difference. Production—the total number of correct items reported in the output sequence—was also included. Thus, each of the 44 output lists had a score for each variable. The shape and semantic variables were expected to be negatively correlated with mean Ln IRT. Mean Ln size difference should be positively correlated with IRT. Production was of interest in that shape, semantic transition, and size variables might (or might not) be positively correlated with it, though no specific predictions were made about those relations. Because the results of this analysis varied slightly depending on which measure of IRT was used, I have taken the mean of the mean and median

Table 15

Averaged Correlations Between IRT and Other Variables (N = 44)

Variable	1	2	3	4	5	6	7	8	9
1. IRT[a]	----	**-.35**	**-.30**	**-.33**	**-.23**	**.33**	**-.01**	**.04**	**-.46**
		.02	*.053*	*.03*	*.13*	*.03*	*.77*	*.78*	*.002*
2. *M* ShapeSim		----	**.55**	**.40**	**.23**	**-.30**	**.05**	**-.02**	**-.01**
			.0001	*.007*	*.13*	*.046*	*.74*	*.90*	*.95*
3. ShapeARC			----	**.51**	**.38**	**-.32**	**.07**	**.05**	**.09**
				.0004	*.01*	*.03*	*.63*	*.75*	*.54*
4. *M* Semantic				----	**.87**	**-.53**	**.04**	**.32**	**.24**
					.00[b]	*.0002*	*.82*	*.03*	*.12*
5. SemantARC					----	**-.32**	**.06**	**.32**	**.30**
						.036	*.70*	*.03*	*.051*
6. *M*LnSizeD						----	**-.13**	**-.45**	**-.23**
							.39	*.002*	*.13*
7. *M*LnVerbRL							----	**.29**	**.36**
								.058	*.02*
8. *M*NROutPos								----	**.58**
									.00[c]
9. Production									----

Notes. Pearson correlations are in **bold**; two-tailed *p*-values are in *italics*. The above results involve means of the means and medians of four analyses, each analysis using a different IRT measure.

[a]IRT is the average of four different measures, mean, median, mean Ln, and median Ln IRT. [b]$p = 2.28 \times 10^{-14}$ [c]$p = 3.18 \times 10^{-5}$.

correlation results for four separate measures of IRT. Those four measures included median (1) and mean (2) IRTs, and median Ln (3) and mean Ln (4) IRTs. The results of this averaged correlational analysis are shown in Table 15.

The shape and size variables were correlated significantly in the predicted directions with IRT, confirming perceptual theories. Mean semantic ARC, though in the expected direction for both perceptual and amodal theories, did not reach significance.

Shape similarity, Ln size difference, and semantic transition were significantly correlated with each other. This is consistent with the classification results reported earlier, which showed significant associations between the three variables.

Mean shape similarity and shape ARC were not significantly correlated with production, while mean size difference and mean semantic transition were marginally correlated with production. Only semantic ARC reached borderline significance in its correlation with production. The lack of correlation between mean shape similarity and production could be due to the fact that there were more SemCl—ShapeSw than SemSw—ShapeCl transitions. Semantic clustering in the absence of shape clustering would tend to reduce a correlation between mean shape similarity and production.

Multiple regression analysis was carried out on the mean scores, examining the portions of variance in IRT accounted for by the variables of interest. Again, the results are averaged over the four analyses, each of which used one of the measures of IRT. Only the specific variables that had the strongest relation with IRT were included in the analysis. Consequently, shape and semantic ARC were removed, and median replaced mean Ln size difference. Portions of variance were measured using partial and part correlations. As Tables 16 and 17 show, of the three main variables, median Ln size

Table 16

Averaged Regression Coefficients and Significance Tests (N = 44)

Component	*B*	SE_B	*β*	*t*	*p*
(Constant)	-4.007	2.456		-1.24	.14
M Semantic	-0.611	0.879	-.07	-0.60	.29
M Shape Sim	-0.339	0.190	-.20	-1.81	.04
*Mdn*Ln SizeD	1.575	0.552	.36	2.86	.004
M Ln VerbRL	0.570	0.518	.15	1.35	.10
M NROutPos	25.569	4.737	.67	4.89	.000017
Production	-0.084	0.011	-.80	-6.35	3.3×10^{-7}

Notes. Combined mean of mean and median regression results. *P*-values are one-tailed.

Table 17

Averaged Correlation for Each Variable with IRT (N = 44)

Variable	Zero-Order	Partial	Part
M Semantic	-.33	-.10	-.06
M Avg Shape Similarity	-.35	-.29	-.18
Mdn Ln Size Difference	.33	.43	.28
M Ln Verbal Response L	-.01	.22	.14
M N Rank Output Position	.04	.62	.48
Production	-.46	-.72	-.63

difference on average had the strongest unique association with IRT duration, followed by mean shape similarity and, lastly, by semantic transition.

Using the standard method of multiple regression, the averaged R^2 = .63, adjusted R^2 = .57, and $F(6, 37)$ = 10.92, $MS_{residual}$ = 0.539, $p = 8.3 \times 10^{-7}$. Multicollinearity was not a problem: The mean variance inflation factor (VIF) values in this set ranged from 1.2 for mean Ln verbal response length to 1.9 for mean normalized rank output position.

The fact that the perceptual variables each, on average, accounted for more variance in IRT than semantic transition here could be due to the fact that participants' shape similarity means and size difference medians, respectively, were derived from more precise ordinal measures, as compared to the means based on the coarse dichotomous (0 = switch, 1 = cluster) semantic measure. To examine this issue, another multiple regression analysis was carried out, the same as the one immediately above (N = 44), except that mean dichotomous shape (0 = switch, 1 = cluster) replaced mean shape similarity, mean dichotomous size difference (median split of low vs. high) replaced median Ln size difference, and the original sample from which the output list scores were derived had N = 1301 instead of N = 1387. The mean of the analyses using the four IRT measures was assessed. The dichotomous shape part r = -.10, the semantic part r = -.15, and the dichotomous size difference part r = .17; none of these reached significance. Only normalized rank output position and production were significant in that analysis. Hence, using a coarse dichotomous level of measurement here does result in some loss of information. In addition, one should expect there to be some loss of information due to using, as the units of analysis, mean (or median) scores from whole FE output lists instead of individual inter-item transitions. The size difference median or mean for an FR output list also would not be sensitive to the difference between semantic cluster and switch transitions. (Size difference is associated with IRT durations at semantic cluster transitions but not at semantic switch transitions). The semantic part ***r*** is only -.15, which is less than the part ***r*** for dichotomous size difference, and which is not much higher than the part ***r*** for dichotomous shape. One should expect semantic transition to account for *much* more variance than either shape or size. It is possible that this size difference result is specific to this analysis of mean (or median) scores, which may be insensitive to differences associated with semantic cluster versus switch transitions. This possibility will be examined in the course of the next analysis on the pooled data, presented below.

2.6.6.2. Regression Analysis of the Pooled Data at Lagged Intervals. First, a multiple regression analysis was carried out on the relevant pooled IRT sample (N = 1387) in order to obtain approximations of the coefficients and their standard errors. These values were noted for later comparison. Another variable, Number of IRTs in an output sequence, was added to capture additional variance and to further reduce the autocorrelation.

Next, analyses were carried out to find the inter-IRT lag at which autocorrelation was no longer present. Autocorrelation was measured using the Durbin-Watson statistic. Although the Durbin-Watson statistic only tests for lag-1 autocorrelation, in this analysis SPSS (13.0, Grad Pack) "ignores" the skipped IRTs (omitted from the analysis using the "select if" option) and treats the *i*th IRTs as though they are directly adjacent, that is, lag-1. Thus, the Durbin-Watson test on a sample consisting of every *i*th IRT gives an indication of how much lag-1 autocorrelation is still present in the *i*th interval sample. Although the autocorrelation and partial autocorrelation function plots indicated that autocorrelation in the pooled IRTs was not significantly present beyond lag-4 (with 95% confidence limits), the Durbin-Watson is a stringent test that has a large indeterminate region wherein the presence of autocorrelation can neither be accepted nor rejected. To conclude that autocorrelation is not present, the obtained Durbin-Watson value must be outside the bound of the indeterminate region in the direction that is closest to approximately 2 (see Neter, Kutner, Nachtsheim, & Wasserman, 1996). Through conducting multiple regression analyses at progressively longer distances between *i*th IRTs, I found that using every 8th IRT removed the lag-1 autocorrelation decisively. Arriving at that finding involved carrying out multiple regression on 8 different subsamples (or sets), using every 8th IRT of Set 1, every 8th IRT of Set 2, and so on, up to and including every 8th IRT of Set 8. A Durbin-Watson test result was obtained with the multiple regression analysis for each of the eight samples. The eight Durbin-Watson obtained values were then compared against the corresponding critical values for the 5% significance level (for critical values, see Cummins, 2006). Note that there were approximately 173 IRTs on average per sample (approximately 170 for size difference). It was found that the upper bound of the indeterminate region ($M_{DW:upper} = 1.82$) was less than the obtained values ($M_{DW:obt} = 1.99$), $t(7) = 2.65$, $SE = 0.063$, $p = .03$ (two-tailed). It is safe to conclude that lag-1 autocorrelation was not present in the lag-8 samples.

The above procedure, using the 8 samples, captured the full pool of 1387 IRTs (1358 for size difference). There were some minor differences among the eight samples. Therefore the correlation and regression results below (Tables 18, 19, and 20) are averaged (i.e., taking the mean) of the mean and median results of the 8 samples. (Note that Ln IRT was the best and only IRT measure used in all of these analyses, though other analyses using raw IRTs also showed the same pattern of results). Nevertheless, the decision regarding the statistical significance of the relations among variables was not affected by whether mean or median values were used.

The averaged zero-order correlations amongst the variables of interest are shown in Table 18. Average shape similarity was transformed using an exponent, 2.5, in order to increase (slightly) the amount of explained variance, though the results are significant for shape similarity whether the transformed or original variable is used. Shape similarity$^{2.5}$,

Table 18

Averaged Correlations Between the Variables of Interest (Pooled)

Variable	1	2	3	4	5	6	7
1. Ln IRT	--	**-.42** *.000*[a]	**-.33** *.0005*	**.26** *.001*	**.21** *.008*	**.45** *.000*[b]	**-.12** *.17*
2. Semantic		--	**.36** *.005*	**-.40** *.000*[c]	**-.07** *.40*	**-.00**[d] *.55*	**.06** *.28*
3. ShapeS.$^{2.5}$			--	**-.31** *.001*	**-.07** *.36*	**-.04** *.59*	**.02** *.48*
4. SizeDiff.Ln				--	**.02** *.58*	**-.03** *.65*	**-.05** *.40*
5. VerbalR.Ln					--	**.19** *.046*	**.20** *.02*
6. LnNROutP						--	**.04** *.64*
7. N IRTs							--

Notes. Mean n = 173; for Ln size difference, mean n = 170. Averaged p-values are two-tailed.
[a]$p = 1.26 \times 10^{-6}$, [b]$p = 3.26 \times 10^{-7}$, [c]$p = 2.69 \times 10^{-6}$, [d]r = -.00025.

Ln size difference, and semantic transition were all correlated significantly, in the predicted directions, with each other and with IRT. In addition, verbal response lengths increase as output position increases. The number of IRTs in an output sequence (N IRTs) did not have a significant zero-order correlation with Ln IRT. However, besides helping to reduce autocorrelation, the N IRTs variable was useful in the regression and partial and part correlational analyses. There was also a subtle but significant tendency for verbal response length to increase for higher N IRTs.

The averaged lag-8 results for correlations (zero-order, partial, and part), β_i for the individual variables, and R^2 and adjusted R^2 for the set of variables, approximated closely the results on those measures for the N = 1387 pooled data. Using the standard method of multiple regression, the averaged $R^2 = .46$, adjusted $R^2 = .44$, and $F(6, 163) = 23.16$, $MS_{residual} = 0.393$, $p = 1.26 \times 10^{-18}$. (The average n for that analysis was approximately 170, which was the average n for size difference). The mean VIF values ranged from 1.05 for Ln verbal response length to 1.26 for Ln size difference, indicating that multicollinearity was not a problem. Compared to the results for the original pooled (N = 1387) data, the removal of the autocorrelation and the reduction in (mean) sample size for the lag-8 analysis led to increases in the standard errors of the regression coefficients (SE_B), and thus to increases in the p-values, as should be expected (see Neter et al., 1996).

Table 19 shows the individual β_i and p-values for each of the explanatory variables. Table 20 shows the zero-order, partial, and part correlations for each explanatory variable. Shape similarity$^{2.5}$ was significantly associated with Ln IRT duration, $p = .01$ (one-tailed), $\beta = -.18$, part $r = -.16$. This is only slightly less than the part $r(42) = -.18$ in the analysis of participants' mean scores reported above. In the

Table 19

Averaged Regression Coefficients and Significance Tests (Pooled)

Component	B	SE_B	β	t	p
(Constant)	1.6355	.2401	--	6.90	7.9×10^{-9}
Semantic	-.5099	.1108	-.30	-4.61	.0002
Shape Similarity$^{2.5}$	-.0039	.0014	-.18	-2.83	.01
Ln Size Difference	.1383	.0941	.09	1.39	.11
Ln Verbal R. Length	.2792	.1270	.14	2.27	.032
Ln N R Output Pos.	1.8954	.2599	.43	7.28	6.35×10^{-9}
Number of IRTs	-.0108	.0046	-.14	-2.28	.026

Note. Reported p-values are one-tailed. Mean n per subsample = 170. Total sample N = 1387.

Table 20

Averaged Correlation of Each Variable with Ln IRT (Pooled)

Variable	Zero-Order	Partial	Part
Semantic	-.42	-.34	-.27
Shape Similarity$^{2.5}$	-.33	-.22	-.16
Ln Size Difference	.26	.11	.08
Ln Verbal Response Length	.21	.18	.13
Ln Norm. Rank O Position	.45	.49	.42
Number of IRTs	-.12	-.17	-.13

present analysis, the association between Ln size difference and Ln IRT did not reach significance, $p = .11$ (one-tailed), $\beta = .09$, and part $r = .08$, though it was positive in all 8 of the subsamples. This is substantially less than the result found for median Ln size difference in the participants' mean scores analysis, where part $r(42) = .28$. Semantic transition was significantly associated with Ln IRT durations, $p = .0002$ (one-tailed), $\beta = -.30$, part $r = -.27$. This is obviously much higher than the part $r(42) = -.06$ obtained in the participants' FE output mean scores analysis. Another difference between the mean scores ($N = 44$) analysis and the present one using pooled individual inter-item transitions (based on $N = 1387$) is that the former gives equal weight to each participant's output list, whereas the latter necessarily gives more weight to those output lists that contain more IRTs. Also, again, the mean scores analysis is insensitive to differences between semantic cluster versus switch transitions (e.g., that size difference is related to IRT duration within semantic cluster but not within semantic switch transitions). Those differences could account partly for the discrepancy between the results of the two analyses, which was most dramatic for semantic transition and size difference. The remaining three variables—Ln verbal response length, Ln normalized rank output position, and number of IRTs—were each significantly associated with Ln IRT durations when the correlations with the other variables were taken into account.

To find the unique association between shape similarity and Ln IRT durations within different subsamples of interest, separate multiple correlational analyses, using the above six explanatory variables, were carried out within domains, trials, semantic transition, and first versus second half of normalized output. These analyses were carried out the same way as those reported above for the lag interval pooled ($N = 1387$) IRTs.

Within each domain, shape similarity was uniquely associated with Ln IRT duration, but the association differed in magnitude between domains. In particular, the magnitude of that association was relatively high within the round domain (shape similarity$^{2.5}$ part r = -.25), higher even than for semantic transition (part r = -.21), but was low in the flat domain (shape similarity$^{2.5}$ part r = -.11; semantic part r = -.28). The mean correlations for shape similarity$^{2.5}$, Ln size difference, and semantic transition in straight and long domains did not depart substantially from the averages displayed in Table 20.

The unique relation between shape similarity and Ln IRT was only slightly stronger on the first (part r = -.18) as compared to the second trial (part r = -.15). The unique relation between dichotomous shape (cluster = 1, switch = 0) and Ln IRT was also only slightly stronger on the first (part r = -.16) compared to the second trial (part r = -.13). Semantic transition and Ln size difference were associated with Ln IRT at magnitudes that were relatively consistent between trials.

At semantic cluster transitions, the unique association with Ln IRT for shape similarity$^{2.5}$ was part r = -.19, and for Ln size difference was part r = .16. At semantic switch transitions, the unique association with Ln IRT for shape similarity$^{2.5}$ was part r = -.15, and for the Ln size difference was part r = .04. Thus, for these data, the association between shape similarity$^{2.5}$ and Ln IRT is relatively independent of the semantic variable, while the association between Ln size difference and Ln IRT is dependent on the semantic variable. Similar analyses were conducted in shape switch versus cluster, and low versus high size difference subsamples. These results again showed that semantic transition and Ln size difference were more dependent on each other for their relation with Ln IRT than was the case for shape similarity$^{2.5}$.

In the first half of normalized output, semantic transition was less strongly related (i.e., part $r = -.22$) to Ln IRT than in the second half (part $r = -.39$); the reverse was true for output position (first half, part $r = .43$; second half, part $r = .10$). The N IRTs variable was more strongly related to Ln IRT duration in the second half as compared to the first, but verbal response length was more strongly related to Ln IRT duration in the first as compared to the second half. For shape similarity$^{2.5}$ and Ln size difference, there was very little increase in correlation with Ln IRT in the second half.

Finally, a partial correlation analysis using the whole pool of individual IRTs (N =1358 for size difference) focussed on the relationships between shape similarity$^{2.5}$, Ln size difference, and semantic transition. Controlling for all other variables, the partial correlation between Ln size difference and semantic transition was $r = -.32$; between shape similarity$^{2.5}$ and semantic transition was $r = .25$; and between shape similarity$^{2.5}$ and Ln size difference was $r = -.22$. This provides further confirmation that, for these FE domain outputs, size difference and semantic transition are more interdependent than shape similarity and semantic transition, and the latter are in turn more interdependent than shape similarity and size difference. These results are also consistent with the pattern of zero-order correlations between the shape similarity, size difference, and semantic transition variables, reported above.

2.6.7. Reported Thoughts in Written Answers to the Post-Task Questions

In responding to the questions about how they retrieved objects, participants generally noted multiple types of thought contents or schemes. For convenience, I will call these thoughts "strategies." Although strategy(ies) is the conventional term, there are limitations in applying it broadly to the thoughts described by participants in these FE tasks. Strategy can imply a high level of premeditation and planned, highly controlled execution. That limited meaning does not capture the variety of thought contents and processes reported by participants. I use the term strategy with that caveat.

A report of a strategy (e.g., Imagery) was counted as a single report per question, even if the participant mentioned it more than once for an FE trial. Each trial's question was scored separately, such that if a participant reported a strategy on one trial's question and reported it again for the other trial, that strategy would be counted once for each trial. A given statement could be scored for more than one type of strategy. For example, one participant wrote: "...I would...think of objects in my fridge that are round, mostly food objects..." This was scored as Location (inside fridge) and Semantic categorical (food).

The names and descriptions for the various strategies are presented in Table 21. Participants provided the terms and descriptions in their responses, which were later classified by the researcher after all the answers had been collected and read. The classifications are tied closely in content to the example descriptions provided by participants. In examples below, note that the words inside a quotation are the participant's, including portions in parentheses ( ). I have added words in brackets [ ] inside a quotation to clarify the meaning. Unless noted otherwise, the reported strategy was attributed generally to facilitation of retrieval, not to difficulties encountered.

Table 21

Types of Reported Strategies

Strategy	Description
Location	Refers to the locations thought of in retrieving object names. The most frequent included one's own home, apartment, or living space, but also included parent's home, friend's or relative's home, classroom, campus, workplace, indoor, outdoor, street, mall, downtown, or store. Specific rooms were often mentioned (e.g., kitchen, living room, bedroom, etc.). Less frequently, some objects were reported to have been found inside other objects such as fridge, car, or bookbag—these too were scored as Locations. Note that scoring of Location here did not include the testing room (see below).
Imagery	Explicitly uses terminology such as "pictured," "visualized," "imagined," "imaged," "scanned," "looked," etc., in describing the strategy. Example: "I would flash pictures of my home and the university in my head and kind of look at the picture to find something round." Although other descriptions may have implied imagery (e.g., Location and Mental Travel), a statement had to have an explicit imagery term to be scored as Imagery. When imagery terms were used, they occurred most often with reference to "seeing" objects in some spatial context (i.e., Location), but sometimes also referred to imaging a representation of the shape category or its general boundaries.
Self In Environment	Refers explicitly to imagining oneself placed in the environment(s) where object items are retrieved. Examples: (1) "...pictured myself in those places to see what I could see"; (2) "...imagined myself being there and thought of every possible thing around me"; (3) "I then placed myself in familiar places...and...[would] mentally 'look' around for round things."
Mental Travel	Refers explicitly to imagining moving physically through some environment where object items are retrieved. Examples: (1) "...I then...began 'walking through downtown'..."; (2) "...I would imagine myself...(e.g., flying through a heavily industrialized and populated city) and imagine what I could see there." Although mental travel was often implied, it was only scored if mentioned explicitly.
Shape Similarity	Refers to inter-object shape similarity or shape-based reminding as a basis for retrieving items. The shape similarity is at a level subordinate to the general shape domain (e.g., spheres, within round). Examples: (1) "When I had exhausted that [other strategy], I took what I had

	already listed (i.e., ball bearings) and allowed it to serve as an inspiration for other, similar round objects (i.e., marbles),"; (2) "[for flat objects]...try to think of shapes which are frequently flat (circle includes pancake, shield, looney)..."
Semantic Categorical or Relational	Reports a general semantic category or relation as a basis for retrieving items. This could be indicated by explicit mention of a category or through a clear example of a semantic association. Examples: (1) "...associate other objects from the idea of one. For example, from 'Book' I thought of 'Newspaper'"; (2) "...I thought of things such as Sports, Foods, etc."; (3) "I generally put things into bigger categories, thought of a category, then listed all long things I could within that area (i.e., Sports: golf club, hockey stick, baseball bat, etc.)."
Unspecified Relational	Indicates retrieving items that were related or linked in some way, but does not specify the nature of the association or give a clear example. Examples: (1) "...I could think of more flat objects that are related in some way"; (2) "...I used one object sometimes to connect me to a different object."
Personal Knowledge	Refers explicitly to self or some autobiographical fact in describing the basis for retrieving items. These reports were less contextualized than Location and Episodic, did not refer directly to events or episodes, and were more self-referential than Semantic categorical/relational. Examples: (1) "I would think of things that might be associated with round things. For example, [in thinking of] my dog, I first thought of frisbee"; (2) "...thinking of things that are most common in my life, for example the sports references and the CD."
Episodic-Experiential	Mentions lived events or experiences as a source of item retrieval. Examples: (1) "...things that I see every day..."; (2) "...recalling things that I have recently done..."; (3) "tracing footsteps of my life"; (4) "I remembered different flat objects from my past experiences with them..."; (5) "...thought about my current day and all of the round things I encounter."
Criteria for Object or Shape	These reports were generally associated with difficulties encountered in the course of the task. The person reports having thought of the definition or critical features that could qualify or disqualify items as objects, or as members of the shape domain. For example, a participant wrote (1) "...[if I thought of long] objects with an apex or base that is shaped differently than the remainder of the object, such as 'spatula,' this would cause me to pause." One participant commented that (2) because rope and string could be balled up short, he or she did not report them for long. Another participant, referring to difficulties

encountered in the straight task, wrote (3) "I would think of an object that had…straight, and curved sides, which could not be included." Another noted: (4) "…when I wasn't completely sure if the object was round, I wouldn't say it."

Previously Reported Content — When cited, this generally involved some difficulty in task performance. (A normal part of FE performance, though not mentioned explicitly in the instructions, is having to keep track of, or otherwise having potentially available, what one has already listed). The person reports having thought of objects listed previously within the trial or in a previous trial (practice task or first shape domain). Examples: (1)"The rounded objects task interfered with this one a bit, as I thought of more rounded objects while naming straight ones,"; (2)"I couldn't get previous answers [from within the same trial] out of my head,"; and (3)"I [had difficulties] trying to remember if I had already said it…"

Switched — Explicitly reports having changed from one subject matter or method to another in order to continue to retrieve items. Although switching can normally be assumed to have occurred several times within every trial for every participant—given that multiple categories or contexts were normally mentioned—switching was only scored once, and only if a transition was explicitly described directly or through narration. Examples: (1) "I found it hard to identify many flat things in the same group so I found myself getting frustrated and switching back and forth…"; (2) "Once I ran out of objects I started thinking of different things like my household, sports and equipment, etc."; (3) "Once I had exhausted these options [real locations]…I began to think of [media sources],"; (4) "Originally many objects initially come to mind. After a long list at the beginning I began to imagine an area, empty box, which was long. I tried to imagine long items that would fit into the box. Once I had difficulty with this task I envisioned my house. I would visually walk around my house searching for things which were long."

Objects in Testing Room — Explicitly reports having thought of items in the testing room. This strategy and Location (see above) were scored as if they were mutually exclusive. The strategy, when reported, did not appear to have been pursued very far, given the limited number of relevant objects in the testing room versus the total objects typically reported per trial, and the numerous other strategies typically mentioned by the same participant. Example: "I began with what was immediately in the interview room."

Other-Miscellaneous — Any strategy that happened to be reported only once or twice within this study. Other was scored once per each different additional strategy mentioned. Not all examples are cited for each type. (a) *Unknown*: Reports that items came to mind but does not implicate a strategy, a cause, or a source. Even a general relation per se is not implied. Examples: (1) "…just popped into my head…" and (2) "…things I could think of immediately." (b) *Orientation*: "When I thought of long I would think of tall more often, since those tall objects stand out more. I would be more likely to notice a tree in the forest or in a field than a log." (c) *Media*: " I began to think of things that I had not seen directly, but indirectly through T.V., movies, books (i.e., the Empire State Building)." (d) *Material Similarity*: Wooden. (e) *Physical Connection/ Close Together Spatially*: One object physically connected to another; one object normally positioned close to another, e.g., "…steering wheel and gear shift…" (f) *Directed Word Priming*: "Start with saying the word 'flat' and see what comes to mind. Then say the word 'flat' again and try to complete the object e.g. 'a flat pancake, window, sheet of wood.'" (g) *Imaginal Placement of Objects*: Imaginally placed objects on a desk. (h) *Previous Conversation*: Thought of items related to the topic of a brief conversation that had taken place just prior to the experiment. (i) *Inference*: "I focussed mainly on household items and structure because I knew they had to be straight."

2.6.7.1. Frequencies of Reports For Strategies. Figure 13 shows the frequencies of each reported strategy type for Trial 1 and Trial 2 FE performance. The relative frequencies of the 14 types of reported strategies remained the same from Trial 1 to Trial 2, $r(12) = .96$, $p = 2.99 \times 10^{-8}$ (two-tailed). There was no change in the frequencies of reported strategies from Trial 1 to Trial 2, $t(13) = -0.33$, $SE = 0.430$, $p = .75$ (two-tailed). Remarkably, all of the participants for at least one of the trials reported having thought of objects in their Location contexts. Imagery was reported by 82% of participants for Trial 1, and by 77% for Trial 2. Self-referential and embodied strategies, such as Episodic, Personal, Mental Travel, and Self In Environment, all together, made up a considerable number of reports. Semantic strategies were reported for at least one trial by 64% of the participants. Nine participants (41%) reported having thought of the shape criteria for at least one trial, but only 4 (18%) reported a shape similarity strategy for at least one trial. The possible reasons for the latter result are discussed next, in section 2.7.

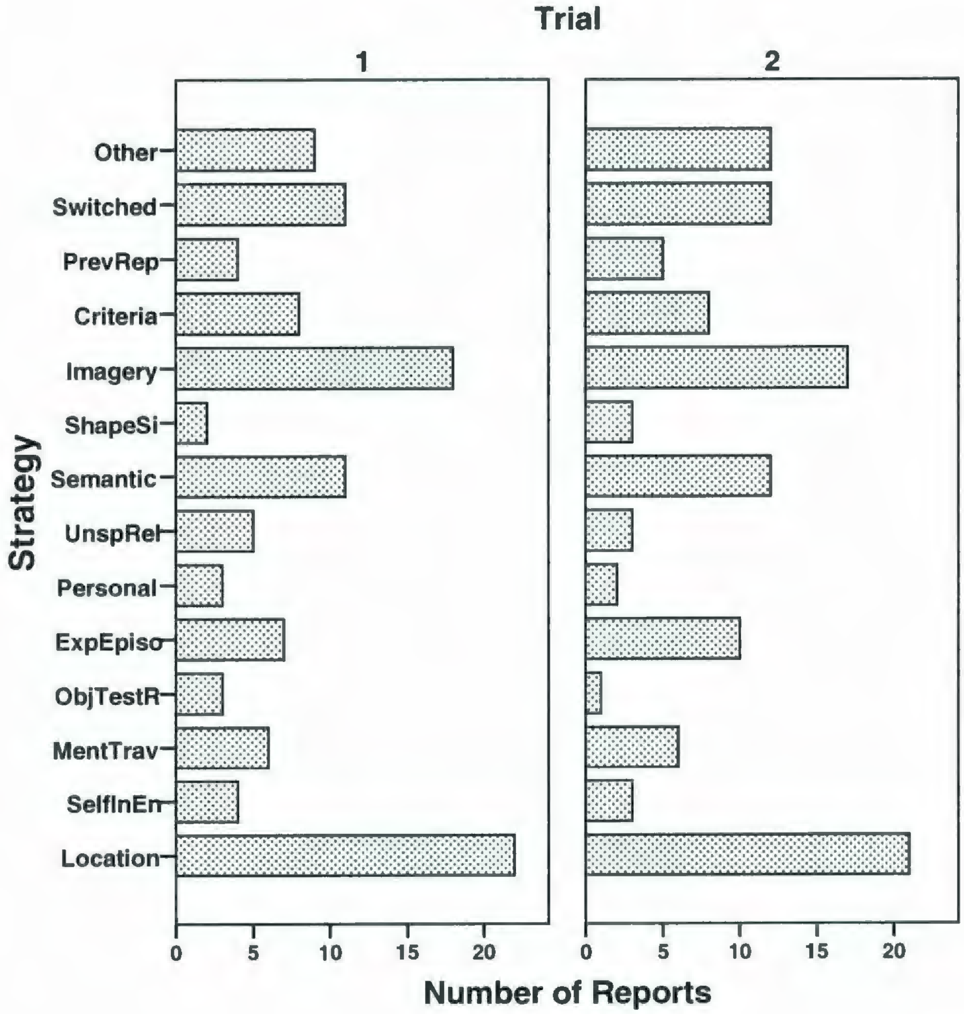

Figure 13. Number of reports of each type of strategy for Trial 1 and Trial 2 FE (maximum = 22). Each type of strategy, if reported, was scored only once per participant, except for *Other* strategies (see Table 21 for details).

2.7. Free Emission Discussion

Consistent with predictions derived from the perceptual theories (Barsalou, 1999; Hebb, 1949), there was significant clustering according to three-dimensional shape subcategories when participants retrieved object names from their LTM or knowledge systems. Shape cluster IRTs were significantly shorter than shape switch IRTs, at both semantic cluster and semantic switch transitions. This pattern of IRT results is generally consistent with the pattern of reaction time results found in lexical decision and naming experiments involving object nouns in the studies by Schreuder et al. (1984) and Flores d'Arcais et al. (1985). Furthermore, IRT durations tended to be shorter when the inter-object shape similarity was higher. This association between shape similarity and IRT occurred across semantic transition type (cluster vs. switch), size difference, verbal response length, and stages of output.

There was significant semantic clustering, consistent with past research. The semantic cluster IRT speed advantage occurred across differences in shape similarity, size difference, verbal response length, stages of output, and FE domains. However, the semantic cluster IRTs were faster at shape cluster as compared to shape switch transitions, and were faster at low as compared to high size difference transitions.

The classification results revealed extensive confounding of object shape, size, and semantic category in peoples' organization of object knowledge. When judges sorted object words into semantic categories, for example, those categories overlapped significantly with the shape and size categories made by other judges on the same object words. Those classification findings were consistent with the compositions of FE participants' output sequences in that transitions of higher shape similarity and lower size difference tended to occur more frequently within the semantic cluster subsample, while transitions of lower shape similarity and higher size difference occurred more frequently within the semantic switch sample. Further evidence of confounding was shown in that semantic transition, shape similarity, and size difference were all significantly correlated with each other, whether the units of analysis were individual inter-item transitions in FE output lists or averaged scores from whole FE output lists. In people's knowledge or LTM systems, then, objects within semantic categories tend to be more alike in shape and size, and objects from different semantic categories tend to be less alike in shape and size. The confounding between object shape and concrete semantic category is consistent with that reported by Rosch et al. (1976), who used other tasks and studied objects that were generally more complex.

The correlation between shape similarity and IRT duration occurred across domains, but was weaker in some domains than in others. For example, in the round domain, shape similarity had a stronger correlation with IRT duration than did semantic transition, whereas, in the flat domain, shape similarity had a weak correlation with IRT duration. Further experiments and more detailed analyses would be needed to find out why the shape similarity-IRT relation is stronger in round than in flat.

The shape cluster speed advantage appeared to be stronger on the first as compared to the second trial. Possibly, there were persistent shape activations from the first trial that interfered with the activations in the different shape domain on the second trial. In future research, this apparent reduction in effect on the second trial could be examined by comparing the shape cluster speed advantage for same-domain (round, round) versus different-domain (round, straight) pairs of trials.

Consistent with the simple perceptual hypothesis that I made before carrying out this study, IRT durations increase as the size difference between successively-listed objects increases. This relation occurs across changes in shape similarity, verbal response length, and stages of output. However, the relation is primarily limited to semantic cluster transitions. Other studies have found that size may be an important dimension within semantic domains. For example, Chan et al. (1993) found that size was a significant dimension in the organization of a subset of mammals from *animals* FE task output. Size appears to be associated significantly with the clustering of mammal names in free recall (FR), though the association was weaker within the birds and fish categories (Caramazza, Hersh, & Torgerson, 1976). Using a variety of tasks, Henley (1969) also found that size emerged as a significant dimension in people's organization of animal concepts, though she did not assess directly the role of size in the FE *animals* clustering obtained in her study. The question remains whether clustering and IRT effects would be shown for size in switching between semantic categories in FE.

Size difference in the present study is a measure of the difference between two ordinal size scores. Possibly, this difference score may incorporate more measurement error than direct ratings of inter-object size difference. Of course, the possible limitations in measurement are unlikely to account for why the association between size difference and IRT duration differed between semantic cluster and switch transitions.

The semantic clustering and IRT effects may be due to other factors. One of those factors is object location. There is a tendency for objects in semantic clusters to be found in the same locations (e.g., kitchen items, bathroom items, etc.). Given the fact that all of the participants in the post-task interview reported having used a search-of-locations strategy to retrieve objects in at least one of the two FE domains, and given that clustering of objects by locations has been demonstrated previously in FE tasks (e.g., see Plumert, 1994; Hirtle & Jonides, 1985), it is probably the case that some significant portions of the IRT and clustering effects classified here as semantic transition are due to the explicit spatial location strategy.

The pervasive use of spatial location/contextual strategies reported here is consistent with findings from a wide variety of FE domains, including fruits, animals, birds, vehicles, furniture, pocket things, and other domains (Rende, 1999, cited in Rende, Ramsberger, & Miyake, 2002; Vallée-Tourangeau, Anthony, & Austin, 1998; Walker & Kintsch, 1985), and names of "people you know" (Bond et al., 1985; Williams & Hollan, 1981; Williams & Santos-Williams, 1980). The pervasiveness of spatial location/contextual strategies probably reflects the adaptation of the brain's neural systems to powerful and ubiquitous real-world constraints. A person needs to be able to remember (or infer) where entities and objects are (or might be) located. FE tasks are in some respects like mental foraging, hunting, gathering, or, in modern contexts, shopping (see the supermarket fluency/FE task, e.g., Tröster et al., 1995). In these various FE tasks, people retrieve numerous objects that are appropriate to some goal or criteria. The ability to travel mentally through various known contexts to search, in advance, for appropriate objects probably saves a great deal of time and effort. One can think through and decide the shortest and perhaps safest travel route to where the most appropriate and/or plentiful objects can be found.

Moreover, people arrange their objects and environments so that things used for the same kinds of activities are in the same locations, thus saving travel and search time.

In their experiences, people tend to interact with similar things in locations that are associated with those types of things. Hence, it is understandable that there could be confounding of location, semantic category, shape, and size in people's output sequences when they retrieve the names of those things in FE.

In their reports of strategies, few participants reported explicit attention to object shape subcategories. Perhaps most people did not consciously engage much in shape subcategorical strategies under these conditions. This is in contrast to the semantic subcategorical organization that probably occurs explicitly in some FE tasks such as *animals*, and to some extent in the shape domain tasks used here. However, it is also possible that people do tend to use shape subcategories, but that in these tasks they may be less aware of shape as opposed to semantic categories. Performing a categorical FE task seems to require that participants consider, at least briefly, the appropriateness of various subcategories as members of the domain, such as reptiles, mammals, birds, and other subcategories within animals; or spheres, disks, cylinders, rings, and so on, within the round domain. Possibly, participants thought of some shape subcategories in order to check that they were listing appropriate objects, but that, in response to the question about how they were able to list so many objects, the most productive strategy—probably, spatial location/contextual search—was reported. Another possible reason for the low frequency of reported shape subcategory strategies is that shape subcategory words may not be as common, relative to semantic category labels, in people's vocabularies, thereby reducing the chances that people would mention shape terms in their post-task reports. Nevertheless, in answering the post-task questions, most FE participants did not use more general terminology that might have implied the use of a shape similarity strategy, such as mentioning that some of the objects looked similar to each other, despite the fact that imagery was commonly reported here. This suggests that physical similarity, including in terms of shape subcategory, was not predominant in people's conscious thinking. Under the present task shape domain restrictions, and due to natural confounding of semantic category with shape, explicit use of a semantic strategy will tend to result in shape clustering. For example, many sports items listed in the round FE domain are spherical. If people generally are more aware of semantic categories as compared to shape categories, then the semantic categories should indeed be more salient in people's actual FE retrieval processes and in their reports.

The open-ended nature of the post-task questions used here has the clear advantage of ensuring that people's reports are not unduly influenced by the experimenter. On the other hand, there are problems with the open-ended format. For example, participants may have used some strategies, such as the shape subcategory strategy, more than was shown in their responses to the open-ended questions. For a variety of reasons, participants may not mention all the strategies that they used. In addition, the procedure does not give an indication of how extensively each strategy was used within a trial. These limitations could be dealt with in future research by giving participants, after the open-ended question about methods of retrieving items, a checklist of strategies that are known from previous research or which are of theoretical interest. Ranking scales could be added beside each listed strategy where participants could indicate how much each strategy was used, how effective it was, and so on.

The mean overall number of items produced by participants within seven minutes for these FE tasks was 36.7, according to the scoring scheme used here. This is close to the mean of approximately 41 items produced in six-minute duration FE of other superordinate domains (professions, clothing, sports, kitchen utensils, four-footed animals, fruit, and crimes), obtained by Graesser and Mandler (1978). If, in future research, some adjustments were made to the present FE task instructions, such as relaxing the object and shape criteria, participants might list a larger number of objects for the shape domains.

The quantities of items produced in the FE domains here are probably less than ideal for the use of multiple regression analyses of item content variables (e.g., shape similarity, size difference, semantic transition) in relation to IRT durations. For example, the part correlation results from the main analysis on the pooled data indicated that only approximately 7%, 3%, and 1% of the variance in IRT duration was accounted for by semantic transition, shape similarity, and size difference, respectively (i.e., altogether, 11%). Yet approximately 18% of the variance in IRT duration was accounted for by the normalized rank output position variable alone. (To my knowledge, it is not yet known what exactly constitutes a high or low amount of variance accounted for in this type of analysis of FE output). To account for substantially more item content-related variance IRT duration, it is probably necessary to use much broader FE domains that contain much larger numbers of objects. For example, multiple regression analyses might show that item content accounts for more variance in IRT durations when analyzing output sequences from domains such as *objects*, *animals*, *foods*, and others that are likely to yield a larger number of responses per unit of time. With steeper deceleration in IRTs, such as in the early phase of FE output or where there is a relatively low number of items and high percentage of between-category switches, there is less variance in IRT explained by item content and more by output position. The much larger numbers of items, and thus the larger number of IRTs produced per person in the broader domains, would better facilitate the use of multiple regression/correlational analyses on individual output lists (see Michela, 1990, for a discussion of within-person correlational analysis).

The proportion of variance in IRT durations accounted for by item content can also be increased, of course, by adding more content variables. Object location would probably account for substantial variance, given its prominence in people's post-task reports of strategies. Other variables of interest could be obtained by having participants label directly their own output transcripts according to their own clustering schemes.

This study was limited in focussing on the relationship between IRTs and the explanatory variables as measured at IRT output position lag 0. Further analyses[13] could take into account lags and leads of 1, 2, and 3, or more, within some reasonable limit. An alternative approach would be to make measurements of IRT and shape similarity, respectively, not just between t_1 and t_2, but also between items t_1 and t_3, t_1 and t_4, and so forth, within a reasonable limit. With that information, for example, the researcher could examine the association between inter-item shape similarity and IRTs measured between the items at those distances. The analysis of inter-item distances (e.g., see Friendly, 1979) could also replace the popular clustering measures, such as ARC, which measure only the directly adjacent (i.e., lag 1 for items) relations between items in the output sequences.

In sum, this FE study revealed significant shape organizational effects in people's retrieval of everyday objects from their LTM or knowledge base. The post-task reports revealed an extensive and pervasive use of search-of-spatial-locations/contexts strategies, suggesting that human memory for everyday objects is strongly organized by object location—a finding that was consistent with past research that used other FE domains. This study also revealed extensive confounding of shape, size, and semantic variables in people's classifications and retrieval outputs of objects. Some of that confounding was untangled statistically in the multiple regression analyses on individual IRTs in FE outputs. However, to arrive at more confident conclusions about the role of object shape in memory organization, it is still necessary to isolate shape from the size, semantic, and other variables, experimentally. The influence of those other variables can be largely removed through careful selection and composition of object lists that are then used in a free recall (FR) study. Hence, the next study set out to examine shape organization under the experimental controls available in the standard FR methodology.

2.8. Study 2: Free Recall (FR) Review

2.8.1. Problems in Previous Recall Studies of Perceptual Clustering

In this review, I examine the methodological problems and issues in previous studies that examined perceptual category clustering in free recall (FR). I discuss how these issues affect our interpretation of the clustering results. Some of the studies mentioned below dealt with perceptual category clustering only incidentally, and were not designed to address the issues raised in this review. Nonetheless, the examination of these issues and problems touches upon general principles that are of concern to anyone who is attempting to assess and interpret the results of any FR study, not just those focussing on perceptual variables.

2.8.2. Organizational Effects in Standard Versus Non-Standard FR

In standard FR, no hints, cues, or instructions concerning organizational strategies are given, and the items are presented in randomized or otherwise unblocked order with respect to the experimenter-defined categories or associations in the list (Kausler, 1974; Murphy & Puff, 1982). Due to the open-ended nature of standard FR, clustering that occurs according to the experimenter-defined categories or associations can be attributed to the natural preexisting organizational tendencies of the participants.

On the other hand, some non-standard recall paradigms that use organizational aids often boost the levels of clustering beyond what is obtained under standard FR conditions. For lists containing semantic categories, clustering in recall is usually increased, relative to the clustering found in standard FR conditions, by instructions to use the categories to aid memory (Gershberg & Shimamura, 1995; Gollin & Sharps, 1988; Miotto et al., 2006), or by blocking the items by category (Cofer, Bruce, &

Reicher, 1966; Gollin & Sharps, 1988; Puff, 1966), or by categorical or property cuing (Cissè & Heth, 1989). Freely sorting items (printed on cards) into groups during study can also lead to higher clustering in recall and increased awareness of the categories, as compared to standard FR encoding of randomly ordered items (Hasselhorn, 1992). In free sorting, participants are not told the categories but are told to organize, and they are able to physically arrange the items in a blocked (non-random) format.

Consistent with the findings for semantic categories, statistically significant increases in secondary clustering have been found for perceptual category lists in conditions with category information or cues (Hudson, 1968, 1969; Wood & Underwood, 1967) and categorical blocking (Wood & Underwood, 1967), as compared to standard FR conditions. Hudson (1968, 1969) did not report whether the low levels of apparently above-chance clustering observed in the standard FR condition were significantly above chance. The standard FR (random-presentation, no-cuing) condition in Wood and Underwood (1967) did not produce above-chance clustering of items by noun referent colour category. Stukuls (1975) found that participants who studied an array of 24 objects that was divided into three framed subsections (i.e., spatial blocking of items) had significantly higher spatial regional clustering in recall than did participants who saw the array of objects without the framed subsections. Clustering by spatial region was significantly above chance in the array-without-subsections (unblocked) condition when participants were given adequate study time (Stukuls, 1975).

Conditions involving categorical cuing or instructions to use categories also appear to increase the difference between cluster- and switch IRT durations, relative to the difference obtained in standard FR, for semantic categories (Kellas, Ashcraft, Johnson, & Needham, 1973; Patterson et al., 1971; Pollio et al., 1969). Because this increase might also occur in the recall of perceptual category lists, it is necessary to first establish whether or not the speed advantage within clusters occurs for the perceptual categories in standard FR.

The above evidence indicates that blocking, cuing, and instructions to organize all normally have the effect of increasing clustering for the experimenter-defined categories beyond what would be obtained on the basis of participants' own unaided organizational tendencies in standard FR. These effects occur for both semantic and perceptual categories. Although organizational aids can be useful methodological tools in the study of memory organization, when the objective of the research is to study people's natural pre-existing organizational tendencies with regard to a particular type of category information, one must examine how those tendencies emerge in standard FR.

2.8.3. Nuisance Associations

Nuisance associations are unintended, unwanted associations and similarities between items within a list. They are present to some extent in all FR lists of the usual lengths, particularly those that contain meaningful items, such as the names of common objects. Items that can be linked by nuisance associations tend to cluster in recall, such that the null hypothesis assumption of chance level clustering in recall is violated to some extent for the target categories of interest.

Wood's and Underwood's (1967) list contained items from different colour categories, but several of the items were also foods [e.g., *rice* (white), *corn* (yellow), *bean* (green), and *beet* (red)]. Participants are likely to become aware of the food items and to cluster them in recall. Food items have several overlapping features and share location and action associations. Consequently, the probability that those items will cluster in colour categories in recall is greatly reduced. These types of nuisance associations would interfere with clustering within perceptual categories such as *smelly, round, soft, white,* and *small* things (Bousfield & Puff, 1964; Hudson, 1968, 1969). Several strong semantic categories of items (food, animal, clothing, etc.), other perceptual categories, and phonologically- and orthographically-similar words, were also included together in the lists used in the studies cited above. It is not surprising, then, that the clustering by the target categories in these studies was, generally, low or minimal.

There were also some within-category nuisance associations. Among Wood's and Underwood's (1967) *green* items, *bean* and *spinach*, and other plant items, would tend to inflate clustering within the target perceptual categories. Thus, some clustering may occur within the perceptual categories that is actually due to other associations.

In preparing lists of meaningful words for use in FR studies, it is practically impossible to remove all nuisance associations. Researchers can minimize nuisance associations by carefully selecting and composing item sets; by using multiple randomizations so that no particular pairings of items occur substantially more often than others; by using more than one item set; and by other methods.

2.8.4. Uncontrolled Interactions Between Different Clustering Schemes

Uncontrolled interactions between category types may have also occurred in the recall of lists composed deliberately by the researchers such that items can be clustered according to more than one experimenter-defined scheme. While this methodology can be useful in comparing the strengths of different schemes, the results are difficult to interpret without first knowing the main effect levels of clustering for each scheme in isolation.

Robertson and Ellis (1987) examined FR of 36 faces (i.e., photos) of famous males from a variety of professions (e.g., actor, athlete, etc.), finding that clustering in recall was above chance for profession, but below chance for face shape categories. The initial finding of below-chance clustering for face shapes was scored based on two very coarse shape categories (round versus thin). To obtain more refined classifications, Robertson and Ellis (1987) then assigned a group of judges to classify the same 36 face stimuli according to face shape categories. This time, using the new classifications for the analysis of the above recall data, the researchers found that the face shape categories did have a "minor role" in recall organization. Yet, the face shape clustering results are not entirely clear because of possible competition between the face and profession schemes.

Uncontrolled facilitation is also possible between alternative schemes included in the same list. Moar (1977) found significant clustering according to shape categories (triangular, square, and circular items) in the standard written FR of the names of line-drawn objects that had been displayed in randomized positions on a study sheet grid. In addition, significant clustering by each of semantic, size, and colour categories was found in recall of other items on other study sheets. Clustering did not reach significance for either phonemic or spatial orientation categories. Significant spatial proximity clustering was also obtained in every study sheet condition. One limitation in this procedure was that the items on a study sheet could have been studied or attended selectively in sequence by obvious category (e.g., shape), which is a systematic rather than randomly-ordered encoding sequence. Hence, the study sheet layout may be used as an external aid to boost organization. A visual presentation of a list of to-be-recalled words, with all the words presented simultaneously for study, can also encourage visual sorting during encoding (Puff, Murphy, & Ferrera, 1977).

Moar (1977) also confounded some of the known features with the depicted features of the objects. For example, large objects (e.g., house) were depicted in 5.1 cm^2 pictures, medium-sized objects (e.g., pram) were depicted in 2.6 cm^2 pictures, and small objects (e.g., cup) were depicted in 1.3 cm^2 pictures, respectively. The size clustering could be due to the depicted sizes, the known sizes, or a mixture of both. Likewise, the shape clustering could be due to the known shapes, the depicted shapes, or both. The same problem of confounding occurred in the colour condition. Moar's (1977) results suggest the presence of perceptual categorical organization in memory, but the confounds between depicted and known features complicate the interpretation of the results.

Frost (1972) and others (Hunt & Love, 1972; Stine, Benham, & Smith, 1987) found greater semantic than depicted-orientation clustering in verbal FR of the names of line-drawn objects. In these studies, each object was drawn in one of four orientation categories: horizontal, vertical, slanted left, or slanted right. These same items also came from four semantic categories: animals, clothing, furniture, and vehicles. The items chosen were very strong and familiar examples of their semantic category—a manipulation that could have reduced the effect for orientation. In addition, the depicted orientations in Frost's (1972) study were arbitrary (e.g., a car drawn on a slant). Hence, the comparison between semantic and orientation clustering cannot be interpreted straightforwardly. Despite these problems, orientation clustering was significant in each of these studies, except for one condition in Frost's (1972) Experiment 2. This condition first engaged participants in practice recall of the names of eight practice pictures that were shown. The problem is that an eight item list is so brief that participants can learn many of the items simply by rote, in the presented order. Serial organization is a normal tendency in FR (Mandler & Dean, 1969). Indeed, there is significantly more serial ordering in FR of such brief lists than (a) would be expected by chance (Nairne, Riegler, & Serra, 1991), and (b) would be found in recall of longer lists (Bousfield & Abramczyk, 1966; Jahnke, 1965). Frost's practice recall task (1972, Exp. 2) may have increased participants' tendency to use serial organization in the 16-item test trial. This could explain why, in the test trial of Experiment 2, orientation *and* semantic clustering were lower for the practice recall group relative to another group that had received a practice orientation recognition test of eight practice pictures.

In an earlier study, Frost (1971) also found significant clustering by depicted orientation[14] category in individually-tested participants' verbal FR of the names of line-drawn objects. In that study, Frost did not include alternative organizational schemes in the stimulus list. Frost's (1971) orientation clustering results appear to be valid. [15]

2.8.5. Within-Category Coherence

In the lists used in previous studies, some categories seemed to be overly broad, containing very diverse exemplars, and thus seemed to lack internal coherence. Bousfield and Puff (1964) classified the words *bracelet, helmet, balloon*, and *platter*, for example, as "round" items, despite the heterogenous three-dimensional shapes of the referent objects. Likewise, *freckles, beet, brick*, and *tongue*, were assigned as "red" (Wood & Underwood, 1967); and *daffodil, sardine, cinnamon*, and *hospital*, were all assigned as "smelly" (Bousfield & Puff, 1964). For similar coherence problems within categories, see Hudson (1968, 1969). It is plausible that the low levels of within-category coherence could have contributed to the fact that clustering for those categories was low or null.

2.8.6. Between-Category Contrast and Alignment

Categorical clustering may be stronger within a high contrast list than within a low contrast list. For example, a list containing rectangular, circular, and triangular categories may yield higher clustering than one containing rectangles of three slightly different length-width ratios. In addition, between-category alignment (e.g., see Hunt & McDaniel, 1993; Medin, Goldstone, & Gentner, 1993) may be important for contrast between categories in FR organization. Contrast may be higher when the categories are all or mostly *aligned* with respect to some common dimension, such as shape in the above examples. However, if a list consists of mixed categories, such as rectangular things, blue things, and heavy things, there is no direct or obvious alignment and no clear contrast between the categories. In past research, lists have been used in which perceptual categories—such as *white, round, soft, smelly,* and *small* (Bousfield & Puff, 1964; Hudson, 1968, 1969)—have been mixed in with semantic and/or ad hoc categories (Einstein & Hunt, 1980; Hodge & Otani, 1996; Hunt & Einstein, 1981; Kroes & Libby, 1971, 1973; McDaniel, Einstein, & Lollis, 1988). Consequently, the categories within those mixed-category lists did not seem alignable in any coherent respect, and perhaps did not contrast with respect to one another. In the conditions where participants were not instructed to organize the items by categories, the lack of alignment may have led to low clustering for those types of mixed lists. Alignment may be especially important for perceptual category clustering according to a single feature, where it is usually the case that for items within the target categories to cluster, participants must switch across numerous other categories.

Despite those limitations, some of the clustering scores were low to moderate, and in some cases were significantly above chance. In the study by Hodge and Otani (1996), participants who were given individual item processing instructions, such as pleasantness rating, or familiarity rating, or individual image generation, had clustering scores that were significantly above chance, with mean ARCs of .220, .214, and .116, respectively. Of course, those mean clustering scores are for the mixed lists; the level of clustering for the sensory-perceptual categories in those lists was not reported.

2.8.7. Ambiguous Words

Many words used in secondary perceptual clustering studies were *polysemous* with respect to category membership. For example, items that were included in the round category (see Bousfield & Puff, 1964; also see Hudson, 1968, 1969), such as *belly*, *balloon*, and *bracelet*, each happen to vary widely in shape, depending on the exemplar and the context. *Belly* may also have been construed as a member of the *soft* category in Bousfield and Puff's (1964) list. Wood and Underwood (1967) included *bracelet* in the colour category yellow, though bracelets occur in a variety of colours; gold, silver, or other metallic colours are probably more common than yellow. There was also *homophony*, with great differences in meaning, such that some nouns did not even refer exclusively to objects: Was *derby* understood to be a black hat, as Wood and Underwood (1967) intended, or was it understood to be a sporting event? In the studies cited above, instructions referred only to "words" to be recalled, with no mention that the words were intended to refer to objects. A test of perceptual clustering by the experimenter-defined categories requires that participants activate the appropriate object representations. The presence of lexical ambiguity makes interpreting the clustering scores problematic, because participants may represent the items in ways that the experimenter did not.

2.8.8. Dominance and Typicality Levels

Exemplar dominance and typicality levels tend to correlate with levels of clustering in FR. *Dominance*, or taxonomic frequency, usually refers to the percentage of participants in a sample who give a specific word as a response to a presented noun. To the stimulus category *animal*, *dog* is high, and *emu* is low, in response dominance. Bousfield, Cohen, and Whitmarsh (1958) showed that, for semantic categories, nouns of high dominance cluster significantly more than do those of low dominance. Bousfield and Puff (1964) found significant clustering for high but not low dominant *smelly* items [e.g., *manure* (high), *hospital* (low)]. Low perceptual clustering in some studies (Bousfield & Puff, 1964; Hudson, 1968, 1969; Kroes & Libby, 1971, 1973; Wood & Underwood, 1967) may have been due to the low-to-moderately dominant items used. These items were obtained from one normative set produced by participants who were instructed explicitly to generate a "sensory" word associate for each noun stimulus (Underwood & Richardson, 1956). A procedure that required participants to produce sensory associates would probably elicit more frequent sensory associates than would the standard word association procedure. However, standard word association norms are more relevant to the standard FR procedure, which also does not require people to think of sensory properties. The dominance levels obtained from the sensory associate procedure are probably overestimated as compared to standard word association and standard FR.

Typicality can refer to how well an exemplar represents its category or property type, as rated on an ordinal scale by participants. Typicality can also refer to the number of features that an exemplar tends to share with other members of the same category as assessed from participants' feature listing, with higher typicality of an exemplar indicated by that exemplar's greater feature overlap with other category members. For example, using either a rating scale or a feature overlap measure, *robin* would generally be considered by most people to be more typical than *penguin* for the category *bird*. For the category *round*, *wheel* may be rated high, whereas *bracelet* may be rated low, on typicality (Katz, 1981, 1983). Exemplars that are high in typicality cluster more than those that are low in typicality in FR of lists containing semantic categories (Bjorklund, 1988; Cissè & Heth, 1989; Hasselhorn, 1992; Kahana & Wingfield, 2000). Typicality may also influence levels of perceptual category clustering. Although typicality ratings were, apparently, not consulted in selecting items for previous perceptual category clustering studies, it is possible that low typicality may have been one of the factors that contributed to reduced levels of clustering in any of the perceptual category clustering studies mentioned in this review.

2.8.9. Adequate Recall is a Prerequisite for Valid Clustering Assessment

Adequacy of recall is a major concern in assessing categorical clustering results (Murphy & Puff, 1982). Obviously, a cluster for any given category cannot occur unless at least two items from that category are reported in recall output. I have argued that a valid assessment of clustering requires a minimum average of approximately two or more items recalled per category (for details, see Appendix G). In Kroes' and Libby's (1973) study of children's FR of a mixed list containing semantic and perceptual categories, where the perceptual categories were round things and white things, the overall mean number of items recalled per perceptual category was less than two. Other recall studies have assessed young children's clustering by shape (Melkman, Tversky, & Baratz, 1981) and colour categories of objects or pictures (Melkman et al., 1981; Melkman & Deutsch, 1977; Perlmutter & Ricks, 1979; Sodian, Schneider, & Perlmutter, 1986). As is common with young children, recall in these studies was very low—in many conditions below an average of two items per category—and categorical clustering was usually quite low for the younger groups (i.e., children in the 3 to 6 years of age range). Caution must be maintained in drawing conclusions about categorical clustering in these cases.

In testing undergraduates, Zimmer (1989, Experiments 1 and 3) did not find clear secondary clustering for shape categories in standard written FR of object nouns (for a summary of those results, see Appendix H). His shape categories included spheroid, triangular, rod/post-like, flat rectangular, block-like, ring-like, disks, and string-like. In his Experiment 2, standard written FR of slide photos of objects was used, and that list was presented twice. Zimmer's study lists contained four exemplars per category; there were 8, 6, and 7 shape categories in Experiments 1, 2, and 3, respectively. Applying my recall adequacy criteria of a minimum mean of 2 items per category to Zimmer's lists (Exps. 1-3) gives a minimum recall of 50%. Clustering by shape category was rather low, but above chance level (zero)[16], when standard mean recall was 58% (Experiment 1, Trial 2, standard FR, mean ARC = .13) and 68% (Experiment 2, standard FR, mean ARC = .12), but there was approximately chance clustering (zero) in conditions where mean recall was 51% or under 50%. For Zimmer's eight reported conditions involving the shape lists (see Appendix H), the correlation between the mean shape clustering (ARC) and mean proportion recalled scores was $r(6) = .90$, $p = .003$ (two-tailed). This raises the possibility that Zimmer's (1989) results for shape may be partly accounted for by the *mean* levels of recall in each condition. However, the issue is not clear because, for Zimmer's (1989) Experiment 1, *individual persons'* shape clustering (ARC) were not significantly correlated with their recall scores. In any case, obtaining clearly adequate recall should give clearer results.

2.8.10. Summary of Methodological Issues in Previous Studies

Problems in previous recall studies that limited the interpretation of the perceptual clustering results included (a) numerous strong nuisance associations; (b) possible interaction effects (competition or facilitation) between different clustering schemes within the same list; (c) low within-category coherence, (d) lack of contrast and alignment between categories; (e) ambiguous words, with uncertain category membership for many items; (f) low exemplar dominance and probably low typicality; and (g) inadequate levels of recall. Any one of these problems could disrupt or obscure a true perceptual clustering effect, and many of the studies reviewed had multiple problems. Despite these challenges, there were some suggestions of moderate perceptual category clustering, including by depicted orientation (Frost, 1971) and shape (Moar, 1977). Reducing the various problems should help to provide clearer tests of clustering.

2.9. Study 2: Free Recall (FR) of Noun Lists Containing Object Shape Categories

2.9.1. Overview and Predictions

In this study, each individually-tested participant heard and later recalled freely a randomized list of object nouns that could be clustered by four shape categories. Standard FR instructions were used (i.e., with no cuing, no instructions to organize, no mention of any categories, etc.). After a two-minute distracter task, the participant recalled aloud as many of the object names as he or she could recall. Each participant completed four of these study-distracter-recall trials. After the end of the fourth trial, the researcher interviewed the participant regarding memory processes and organizational schemes used during the FR task.

To assess the generality of any shape organizational effects arising in recall, three different lists of objects were used. Four levels of restricted-random orders were used for the three lists, with each level containing less than the chance level of same-shape pairs. One level had zero (0) same-shape pairs and the others had 1, 2, or 3 same-shape pairs, respectively. The design of this experiment included the within-subjects factor Trials (4) and the between-subjects factors Lists (3) and Number of Same-Shape Pairs in the Study List (4). Interactions between these variables were not of theoretical interest.

Perceptual theories were used to make several predictions. When recall levels are adequate, people should show significantly above-chance mean shape clustering. Mean shape cluster IRTs should be significantly faster than mean switch IRTs. In response to open-ended questions in a post-experiment interview, people should report that their thought and memory processes in the experiment involved imagery (i.e., suggesting activation of perceptual areas of the brain). People's reports of organizational schemes other than shape should include other perceptual similarities, spatial and contextual relations, action relations, and experiential or episodic schemes. As would be expected for any type of clustering, those people who happen to become aware of the shape similarities between items should show higher shape clustering than those who are unaware or partially aware of the shape similarities. Because all of these results had been obtained in a previous FR pilot study ($n = 6$; Mattless, 2003, unpublished), the researcher had reason to expect that they would be replicated in the present study.

2.10. Methods

2.10.1. Participants

Thirty-five Memorial University students took part in this FR experiment. The majority were enrolled in a first-year psychology course and the remainder were undergraduate psychology students. Data from 2 participants were excluded due to a combination of problems[17], leaving a total of 33 (17 females, 16 males; median age = 18). All participants were volunteers who were paid for their time, and were treated in accordance with the ethical guidelines as mentioned in the FE study. None of the participants had done FR or FE tasks before, and none had heard the critical information about objects, shapes, clustering, and so on.[18]

2.10.2. FR Conditions

Each individually-tested person completed four study-distracter-recall trials of a single list of 20 object names. The list items were presented in different random-restricted sequences on each trial. Three different object word lists were crossed by four different numbers of same-shape pairs in the study list. Lists and numbers of same-shape pairs were between-subjects factors. Each person was assigned randomly to list and number of same-shape pairs groups, with the constraint that there would be approximately equal numbers of people, and of males and females, per group.

2.10.3. Setting

The testing room and its contents were similar to those described previously for the FE study. The experimenter and participant were seated at the table at a 90° angle to each other, with the participant to the left of the experimenter.

Two of the participants each linked two or three items in memory as "things in the room." Therefore, these items, from an initial version of List 3, were removed and replaced for the main version of List 3 presented to subsequent participants (see Appendix I for the Note under Table I3 for List 3). For List 1, *wire* was present for the clip-on microphone and for the audio recorders. However, no participant in this study mentioned the presence of the wires having affected their recall or organization.

2.10.4. Lists

Three different lists were composed. Each list had five exemplars in each of four shape categories (see Appendix I). The majority of the object names were obtained from the previous FE study and from other previous pilot work. The four categories in each list were of simple three-dimensional shapes, as follows. *List 1*: spheres, string-like things, long pointed cylinders, and flat rectangles (not long); *List 2*: cones, long cylinders (not pointed), block-like rectangular forms, and disks; *List 3*: rings, flat rectangles (slightly long), string- or strip-like things, and medium-long cylinders. Using three different lists provided an opportunity to gauge the generality of the results. No specific differences were predicted between lists, categories, or exemplars used.

The items had the following characteristics: moderately high mean typicality; low exemplar variability with respect to shape and low lexical ambiguity; a name that was not simply a shape name (e.g., not simply ball, cone, etc.); low mean shape word association strength (see section 2.11.10 for details); not difficult to pronounce; and familiar to participants. Assessments of item characteristics, such as shape typicality, word length, and object size, are presented in Appendix J.

The following considerations guided the process of composing the lists:

1. Nuisance associations, especially those within shape categories, should be minimized.

(Location, activity, and semantic associations, and strong word associations were especially avoided). 2. Items within a category should be reasonably similar but not identical to each other in shape. 3. Shape categories should contrast with each other on at least one major shape feature (e.g., curvature/straightness, elongation/shortness, thickness/flatness); overlap of shape features between categories should be minimized. 4. Mean item characteristics (e.g., word association strengths, typicality, size, word length, etc.) should not differ systematically between categories.

2.10.5. Same-Shape Pairs and Restricted Randomization of Study List

To assess the possible influence of the study list same-shape pairs on shape clustering in recall, four different between-subjects levels of same-shape pairs were used, including either 0, 1, 2, or 3 pairings of same-shape items in the study list. These numbers of same-shape pairings are *less* than the chance level, which is four (4), in a list having 20 items with 5 exemplars in each of 4 categories.

To reduce the influence of nuisance associations and to ensure that items and categories were not confounded with position in the lists, over 100 different restricted-random orders for these lists were used in this study. Each participant received a different restricted-random order of items in the study list for each trial, though the number of same-shape pairs in the study list was held constant across trials for each participant.

The list randomizations involved numerous restrictions. No same-shape item pairing was permitted within the first three or last three items of a list. In the study lists containing more than one same-shape pairing, no same-shape pairings were ever joined to each other to form a larger cluster, and there were generally several items intervening between one same-shape pairing and another. Also, there was never more than one pairing of specific same-shape items per shape category in a study list. Shape categories were also counterbalanced as much as possible with respect to the order in which their exemplar pairings occurred, and the frequency with which they occurred, within and across participants' study list sets. The recurrence of particular same-shape pairings was not permitted within the set of four trials for the same participant, and such pairings were counterbalanced as much as possible between participants.

For a given participant's set of four study list randomizations, any obvious recurrences between trials of specific item pairings were rearranged to avoid the repetition. In addition, any obvious recurrence between trials of an item in a specific list position was moved to a different position to avoid the repetition. The recurrence from one trial's sequence to the next of any item in the first or last position of a list was not permitted.

Many of these restrictions, along with others mentioned under Lists, make the test of clustering by shape a conservative one. Given the natural tendency for people to organize some items in recall based on inter-item proximity or adjacency in the study list (e.g., Kahana, 1996; Mandler & Dean, 1969), particularly those items from the early part of the list, shape clustering in this study should tend to be slightly below chance if it is true that shape clustering does not occur.

2.10.6. Procedures for FR and the Distracter Interval

Each individually-tested participant completed four trials of standard FR. For each trial, following instructions, the participant heard the list, performed a distracter task, and then recalled the list. Standard FR instructions given immediately prior to Trial 1 list presentation informed the participants that they would hear a list of the names of objects, and that they would later be asked to recall as many of the object names as they could (for the detailed instructions, see Appendix K). To reduce perceptual distractions and nuisance mnemonics involving the spatial layout of the testing room, participants were also asked to close their eyes during the study and recall phases. In presenting the list, the experimenter read aloud the object names at a rate of 5 s between word onsets (i.e., the same rate as used in Zimmer, 1989, Experiment 1). During a 2-minute interval between study and recall phases, the participant, with eyes open, counted aloud backwards by threes, continually, from a different large number on each trial (some participants counted back from 400, 600, 800, 1000; others counted back from 300, 500, 700, 900). After this, the participant was instructed to recall as many of the object names as possible. The spoken recall was recorded using the same equipment as was used in the previous FE study. Each of the four trials followed this same format.

The Trial 1 instructions asked participants to "think of" each named object as it was being read. Testing the hypotheses regarding a shape clustering effect requires the activation of object representations. Thus, participants should think of items as objects, not as actions or activities (e.g., *nail, volleyball*); and they should not merely focus on phonology or other surface properties of the words.

2.10.7. Post-Experiment Interview.

Interview questions followed the fourth recall trial. The interview was divided into three stages. These included (1) responses to open-ended questions about thought and memory processes, (2) sorting the cards that bore the object names into remembered groups, and (3) specific questions about awareness of the shape similarities among the objects in the list.

2.10.7.1. Thought and Memory Processes During the FR Experiment. The following open-ended questions were asked to elicit spoken answers:

> 1. Describe what went through your mind while you were hearing the objects' names being presented.
> 2. Describe what went through your mind while you were recalling the objects' names.
> 3. Your recall of the list improved over trials. Besides the benefits of practice and hearing the list more than once, to what do you attribute your improvement over trials?

Answers were obtained for each question before the subsequent question was asked. For vague answers, such as "Some objects were related in my mind," the experimenter sought clarification by asking a follow-up question (e.g., "In what way were they related?"). The experimenter did not audio record this stage or subsequent stages of the interview, but instead took notes while the participant answered.

2.10.7.2. Sorting the Items into Groups and Describing Them. In this task, the participant was shown a deck of cards bearing the names of the 20 objects from the list that he or she had recalled and was asked to sort the cards into groups according to how the items went together in memory during the experiment. The deck of cards was not present in the testing room until the experimenter brought it in for this task. For each participant, the deck had already been reshuffled thoroughly into random order. In introducing the task, the experimenter spread the cards around in a loose, disorderly pile on the table and then began the instructions:

> These are the names of all the objects from the list. The task is to sort them into groups according to how they went together in your memory during the experiment. Any items that didn't seem to go together with other items in your

> memory can be put aside as miscellaneous. Let me know if you have any questions, and please let me know when you are finished.

Once the sorting was finished, the experimenter asked: "Can you describe to me the nature of your groupings, how the objects in each of your groupings went together in your memory? Start with any group you wish." The experimenter then diagrammed the groupings and took notes on the descriptions and labels provided by the participant.

2.10.7.3. Questions About Awareness of the Shape Categories. If the participant did not report awareness of the experimenter-defined shape similarities or categories in the previous questions and sorting tasks, the following question was asked. Q1. "Did you notice during the experiment that all the objects could be grouped together in memory according to different shape categories?" If the participant answered "yes" to Question 1, or had already mentioned, prior to Question 1, an awareness of the shape similarities, then Question 2 was asked. Q2. "Around what phase in the experiment did you notice the shape similarities?" If the answer to Question 1 was "no," then Question 3 was asked. Q3. "Did you notice some shape similarities between some items during the experiment?" If the answer to Question 3 was "yes," then Question 2 was asked. The person's session finished with debriefing.

2.10.8. Measurement and Scoring

After the end of an individual's session, the experimenter transcribed the recorded output sequences for the purpose of scoring and measurement. Recall, clustering, subjective organization, IRTs, qualitative responses to post-task interview questions, and sorted groupings, were all assessed. The scoring of the post-experiment interview responses is described later in the Results, section 2.11.6.

2.10.8.1. Recall. An item was considered to have been recalled correctly where it was reported for the first time within a given recall trial; a later repeat of the item was not counted. Repeats made up 3.7% of the total items reported. A word substitution was counted as a correct recall if the referent was sufficiently similar to the target item in shape and meaning. For example, "coffin" was judged as an acceptable substitute for casket, as was "checkerboard" for chessboard, "pen" for pencil, and so on. These acceptable substitutions made up only 0.9% of the total recall, and most (65%) of them were produced by only one participant. Reported items that were not clear substitutes for an item from the study list were treated as errors. Such errors, called intrusions, were rare (0.5% of the total recall), and were either misinterpretations or associates of list items.

2.10.8.2. Clustering. Two different measures of categorical clustering were used: the Adjusted Ratio of Clustering (ARC; Roenker et al., 1971), and the Modified Ratio of Repetition (MRR; reviewed in Murphy, 1979). The ARC ranges from a maximum of +1 to a minimum of –1, and has a chance level of 0 (zero). The MRR ranges from a maximum of 1 to a minimum of 0 (zero), with a chance level that varies. Formulae for the ARC and MRR are presented in Appendix F. Discussions of these widely-used measures, and justifications for preferring them over a variety of others, are provided in Murphy (1979) and Murphy and Puff (1982).

A categorical repetition was scored when two consecutive correctly recalled items were reported from the same experimenter-defined shape category. Problematic items such as repeats and errors were treated differently in scoring clustering according to ARC and MRR schemes. For the MRR, problematic items were simply excluded from the analysis. For the ARC, problematic items were included to account mathematically for the positions that they occupied, but were not scored as clusters (for further details, see the FE study, section 2.5.5.2.).

The MRR chance estimates for low levels of recall (5, 6, and 7 items recalled) were found to be inaccurate. After carrying out numerous randomizations of lists of those lengths, I found that the MRR chance formula was not accurate for those low levels of recall. I therefore used the means based on numerous actual randomizations as the chance estimates for those low levels of recall (see Appendix L for details).

2.10.8.3. Subjective Organization. Pellegrino's (1971) bidirectional pairwise measure of subjective organization (Tulving, 1962), sometimes labelled as ARC', was used to assess the level of recurrence of individual pairings of items from one recall output trial to the next. This measure does not discriminate between pairings of items in terms of categorical content. Rather, it captures the recurrence of any pairings of individual items on output Trial t_{+1} that were paired on output Trial t. I will refer to this measure as SO-ARC. Justifications for choosing this measure over others can be found in Murphy and Puff (1982). The SO-ARC ranges from a low of –1 to a high of +1 and has a chance value of zero. The formula is shown in Appendix M.

An intertrial repetition was scored for each pairing of items that occurred together on output Trial t and output Trial t_{+1}. For example, if the item sequence *abcdef* occurred in Trial 1 recall output, and *ghfedbci* occurred in Trial 2 recall output, three bi-directional pairings (*fe*, *ed*, and *bc*) recurred. Problematic items such as repeats and errors, when they interrupted a recall trial output series, were included as regular items in the scoring scheme used here for the SO-ARC. Excluding these items, in my judgment, would have led to some dubious scores, particularly for trials having low recall.

2.10.8.4. Interresponse Times (IRTs). IRTs were measured, at the millisecond level, from the onset of one word to the onset of the next word in the output sequence. Wavesurfer (Sjolander & Beskow, 2004) software was used in making these measurements, as described for the FE study previously.

2.11. Results

2.11.1. A Note on the Design and Analyses

The design structure of this study would normally suggest the use of a three-way (3 x 4 x 4) ANOVA on the between-subjects factors Lists (3) and Number of Same-Shape Pairs in the Study List (4), and the within-subjects factor Trials (4). However, that was not the intention of including the various conditions. Were such an analysis to be conducted with this $N = 33$ sample, there would be a mean of only 2.75 participants per cell (i.e., 33 / 12), which does not give a sufficient basis for drawing conclusions from the test of the three-way interaction. It was feasible to analyze the data in separate ANOVAs, one of them on Trials x Lists, and the other on Trials x Number of Same-Shape Pairs. These two-way ANOVAs were carried out on scores for recall, ARC clustering, and subjective organization (SO-ARC), respectively. No significant interactions were obtained. Therefore, in the Results subsections below for recall, clustering, and subjective organization, the results for three separate one-way ANOVAs are presented, providing information about the main effects.

2.11.2. Recall

The number of items recalled increased with successive trials (see Table 22), $F(3, 96) = 154.63$, $MSE = 1.884$, $p = 1.25 \times 10^{-36}$, $\eta^2 = .83$. The number of same-shape pairs in the study list was not associated with mean recall, $F(3, 29) = 0.49$, $MSE = 6.288$, $p = .69$. The difference between the lists on participants' mean recall approached significance, $F(2, 30) = 2.92$, $MSE = 5.347$, $p = .07$. Specifically, the List 2 group had higher mean recall ($M = 14.33$) than the List 1 group ($M = 11.98$), a difference that was nearly significant according to the Tukey HSD post hoc test ($q = 3.35$, $p = .055$).

2.11.3. Clustering

The mean shape clustering ARC score for each trial is shown in Table 22. Above-chance shape clustering was nearly significant on Trial 1, $t(32) = 1.647$, $SD = 0.283$, $p = .0547$ (one-tailed)[19], $\eta^2 = .08$, and significant on subsequent trials [Trial 2, $t(32) = 3.81$, $SD = 0.264$, $p = .0003$ (one-tailed), $\eta^2 = .31$; Trial 3, $t(32) = 3.86$, $SD = 0.232$, $p = .0003$ (one-tailed), $\eta^2 = .32$; Trial 4, $t(32) = 3.68$, $SD = 0.344$, $p = .0004$ (one-tailed), $\eta^2 = .30$]. Similar results were found in comparing MRR obtained versus chance clustering [Trial 1 $t(32) = 1.50$, $SD = 0.275$, $p = .072$ (one-tailed), $\eta^2 = .07$; Trial 2 $t(32) = 3.32$, $SD = 0.201$, $p = .001$ (one-tailed), $\eta^2 = .26$; Trial 3 $t(32) = 3.28$, $SD = 0.178$, $p = .001$ (one-tailed), $\eta^2 = .25$; Trial 4 $t(32) = 3.46$, $SD = 0.260$, $p = .0008$ (one-tailed), $\eta^2 = .27$]. There was close agreement between ARC and MRR obtained scores, despite the different scoring schemes [ranging from $r(31) = .956$ for Trial 1, to $r(31) = .998$ for Trial 4]. Likewise, the correlations between ARC and MRR difference scores (i.e., MRR obtained minus MRR chance) ranged from .952 to .994.

Because the Trial 1 clustering results were marginally significant, a further analysis was conducted for clarification. To ensure that recall was adequate for the

Table 22

Overall Mean Recall and Mean Clustering (N = 33)

	Trial			
Measure	1	2	3	4
Recall	**9.3**	**13.0**	**14.8**	**16.1**
SD	2.85	2.62	3.10	2.23
ARC	**.12**	**.18**	**.16**	**.22**
SD	0.428	0.264	0.232	0.344
MRR obtained	**.41**	**.40**	**.38**	**.42**
SD	0.276	0.197	0.177	0.257
MRR chance	**.34**	**.28**	**.27**	**.26**
SD	0.057	0.021	0.022	0.012

Notes. Maximum possible recall = 20 items. Chance ARC = 0, and the maximum for ARC and MRR = 1. Zimmer's (1989) results are shown in Appendix H for comparison.

purposes of assessing clustering (Appendix G), this analysis was restricted to include those participants (n = 21 out of 33) who correctly recalled eight or more items. For this n = 21 subsample, Trial 1 ARC scores (M = .15) were significantly above chance, $t(20)$ = 1.93, SD = 0.367, p = .034 (one-tailed), η^2 = .10. Using the MRR measure for the same analysis on this n = 21 subsample, obtained clustering (M = .4006) was significantly above chance expected clustering (M = .3054), $t(20)$ = 1.732, SD = 0.25174, p = .0495 (one-tailed), η^2 = .09. Therefore, when there are adequate numbers of items recalled, significant mean shape clustering does occur on the first trial. Note that all of the 33 participants recalled at least eight items on Trials 2, 3, and 4.

The increase in ARC scores over trials was not significant, $F(3, 96) = 0.78$, $MSE = 0.071$, $p = .51$. There was no difference for mean ARC scores between lists [$F(2, 30) = 0.44$, $MSE = 0.057$, $p = .65$]. Mean clustering scores were significantly, or nearly significantly, above chance within each list according to the ARC [List 1 $M = .23$, $t(9) = 2.92$, $SD = 0.251$, $p = .0085$ (one-tailed), $\eta^2 = .49$; List 2 $M = .16$, $t(12) = 2.62$, $SD = 0.220$, $p = .011$ (one-tailed), $\eta^2 = .36$; List 3 $M = .14$, $t(9) = 1.72$, $SD = 0.249$, $p = .059$ (one-tailed), $\eta^2 = .25$] and the MRR [List 1 $M_{obtained} = .45$, $M_{chance} = .31$, $t(9) = 2.46$, $SD = 0.187$, $p = .018$ (one-tailed), $\eta^2 = .40$; List 2 $M_{obtained} = .38$, $M_{chance} = .28$, $t(12) = 2.23$, $SD = 0.162$, $p = .023$ (one-tailed), $\eta^2 = .29$; List 3 $M_{obtained} = .39$, $M_{chance} = .29$, $t(9) = 1.78$, $SD = 0.178$, $p = .054$ (one-tailed), $\eta^2 = .26$]. There were no significant differences in ARC between numbers of same-shape pairs in the study list [$F(3, 29) = 1.17$, $MSE = 0.054$, $p = .34$]. The mean ARC scores for the four same-shape pairs groups (0, 1, 2, and 3 pairs) were as follows: 0 Pairs $M = .21$ ($n = 9$); 1 Pair $M = .27$ ($n = 9$); 2 Pairs $M = .12$ ($n = 7$); 3 Pairs $M = .07$ ($n = 8$). Therefore, the significant clustering in recall cannot be attributed to carry-over of same-shape pairs from the study list.

Although clustering did not increase significantly with trials, the *percentage of participants* who had clustering scores above chance did increase substantially from Trial 1 to Trial 2 (see Table 23). Approximately 51.5% of participants had ARC scores above chance on Trial 1, increasing to 73% on Trial 2, and levelling off at a mean of approximately 71% for Trials 3 and 4. For Trial 1, when the sample was restricted to include only those participants having eight or more items recalled, there was also an

Table 23

Percentages of Persons With Clustering Scores Above, At, or Below Chance (0)

		Percentage of Clustering Scores		
Trial	Measure	Above 0	At 0	Below 0
1[a]	ARC	51.5	12.1	36.4
1[a]	MRR_{diff}	54.5	----	45.5
2[b]	ARC	72.7	3.0	24.2
2[b]	MRR_{diff}	75.8	----	24.2
1[c]	ARC	57.1	9.5	33.3
1[c]	MRR_{diff}	61.9	----	38.1

Notes. MRR_{diff} refers to MRR obtained – chance. Thus, chance $MRR_{diff} = 0$.
[a]$N = 33$. [b]The Trial 2 ($N = 33$) results were similar to those of Trials 3 and 4. [c]This Trial 1 sample was restricted to cases in the present study having eight or more items recalled ($n = 21$).

increase in the percentage having above-chance shape clustering. The fact that approximately 36% of participants had less-than-chance ARC scores on Trial 1 suggests that they used organizational schemes other than the experimenter-defined shape categories. Inspection of participants' first trial outputs suggested a primacy effect, with a tendency for items presented in the initial part of the study list to be recalled together in the initial stage of recall. Other alternative schemes are described later in the Results section (see Post-Experiment Interview).

2.11.4. Subjective Organization

In recalling the list from one trial to the next, participants tended to carry over specific pairings of items more than would be expected to occur by chance. Their subjective organization (SO-ARC) scores were significantly above chance (zero) for the pairing of Trials 1 and 2 [$t(32) = 4.24$, $SE = 0.057$, $p = .00009$ (one-tailed), $\eta^2 = .36$; $M = .24$; with 82% of persons scoring above chance], Trials 2 and 3 [$t(32) = 4.675$, $SE = 0.032$, $p = .000025$ (one-tailed), $\eta^2 = .41$; $M = .15$; with 76% of persons scoring above chance], and Trials 3 and 4 [$t(32) = 4.852$, $SE = 0.038$, $p = .000015$ (one-tailed), $\eta^2 = .42$; $M = .19$; with 82% of persons scoring above chance]. Separate one-way ANOVAs showed that SO-ARC scores did not change significantly for trial pairings [$F(2, 64) = 1.10$, $MSE = 0.062$, $p = .34$], lists [$F(2, 30) = 1.01$, $MSE = 0.021$, $p = .38$], or the numbers of same-shape pairs in the study list [$F(3, 29) = 2.24$, $MSE = 0.019$, $p = .1048$].

2.11.5. Interresponse Times (IRTs)

There were initially 1703 IRTs available for analysis from the 33 participants. IRTs bordered by an item or items that were not scored as correctly recalled, or which contained irrelevant speech, were excluded (7.5% of the 1703 IRTs). IRTs were also excluded from the analyses if they were longer than 15 s (1.5% of the 1703 IRTs). These exclusions resulted in a sample of 1563 IRTs that was used in the analyses below. Due to the high variability and the relatively small number of IRTs per cell, particularly for Trial 1, it was necessary to use mean rather than individual IRTs for the two main analyses. The two main analyses compared mean cluster with mean switch IRTs for *persons* ($n = 33$) and for *items* ($n = 60$). Subsequently, I will examine the IRT type (cluster versus switch) differences over trials and recall output stages, using the whole pool of 1563 individual IRTs.

In the persons analysis, mean cluster IRTs ($M = 2.148$ s) were faster than mean switch IRTs (M = 2.615 s), $t(32) = 2.87$, $SD = 0.937$, $p = .0037$ (one-tailed), $\eta^2 = .20$. [Note that the mean IRTs from the total pool (N = 1563) for cluster and switch were 2.018 s and 2.609 s, respectively]. There were no significant main effects for lists or number of same-shape pairs in the study list, nor were there significant interactions between those variables and IRT type (cluster versus switch). There were insufficient numbers of participants per cell (mean n = 2.75) with which to analyze the Lists x Number of Same-Shape Pairs interaction. The median cluster (M of $Mdns$ = 1.591 s) versus switch (M of $Mdns$ = 1.618 s) IRT comparison was not significant [$t(32) = 0.24$, $SD = 0.641$, $p = .41$ (one-tailed)], but was nearly significant according to the Wilcoxon signed ranks test [$T(33) = 190$, $p = .053$ (one-tailed)].

The use of median scores may be problematic for the analysis by persons because there were often only small numbers of IRTs per participant, whereas the optimal use of medians requires large numbers of scores. Therefore, the median switch versus median cluster IRT comparison was made for object shape categories (n = 12, plus one additional amalgamated category that contained the mean of the median IRTs for bulbs and slab rectangles), where more IRTs were binned per category than per person. IRTs were collected by the category (i) beginning the IRT or (ii) ending the IRT. For the beginning-category analysis, cluster IRTs (M of $Mdns$ = 1.506 s) were marginally significantly shorter than switch IRTs (M of $Mdns$ = 1.670 s), $t(12) = 1.653$, $SD = 0.3569$, $p = .062$ (one-tailed), $\eta^2 = .19$ [Wilcoxon $T(13) = 21$, $p = .043$ (one-tailed)]. The end-category analysis showed that cluster IRTs (M of $Mdns$ = 1.501 s) were significantly shorter than switch IRTs (M of $Mdns$ = 1.663 s), $t(12) = 2.13$, $SD = 0.273$, $p = .027$ (one-tailed), $\eta^2 = .27$ [Wilcoxon $T(13) = 16$, $p = .02$ (one-tailed)]. The difference of approximately 0.162 s between the means of category median cluster versus switch IRTs is considerably larger than the mere 0.027 s difference found in the persons' medians analysis (presented above), and is closer to the difference of 0.186 s for the total pool (N = 1563) median cluster (1.445 s) and switch (1.631 s) IRTs. Note that the effect for category median IRTs is not reduced by removing the amalgamated category from the analyses; removing it increases the difference, to approximately 0.190 s. The effect was larger using category means instead of medians, whether or not the amalgamated category was included.

For the items analyses, the mean cluster IRT and mean switch IRT for each item were used. Because an IRT is bordered by two words, two assessments were made, one in which the IRT was assigned to the word beginning it and the other in which the IRT was assigned to the word ending it. Note that, despite the apparent redundancy, the two scoring methods do not overlap entirely.

For beginning-word IRTs, as predicted, cluster IRTs (M = 2.051 s) were significantly shorter than switch IRTs (M = 2.645 s), $t(59)$ = 4.72, SD = 0.974, p = .000007 (one-tailed), η^2 = .27. (Another analysis confirms these results; see Note[20]). Neither the main effect for list [$F(2, 57)$ = 0.84, MSE = 0.217, p = .44] nor the interaction between list and IRT type [$F(2, 57)$ = 1.19, MSE = 0.471, p = .31] were significant. The cluster speed advantage was significant within each list. For the beginning-word analysis of List 1, cluster M = 2.171 s, switch M = 2.595 s, $t(19)$ = 2.089, SD = 0.9077, p = .0252 (one-tailed), η^2 = .19; for List 2, cluster M = 1.994 s, switch M = 2.486 s, $t(19)$ = 2.43, SD = 0.904, p = .013 (one-tailed), η^2 = .24; and for List 3, cluster M = 1.989 s, switch M = 2.853 s, $t(19)$ = 3.33, SD = 1.133, p = .002 (one-tailed), η^2 = .40.

The end-word analysis results were consistent with those of the beginning-word analyses: Cluster IRTs (M = 2.068 s) were significantly shorter than switch IRTs (M = 2.711 s), $t(59) = 4.44$, $SD = 1.120$, $p = .00002$ (one-tailed), $\eta^2 = .25$. There was no main effect for lists, $F(2, 57) = 1.17$, $MSE = 0.687$, $p = .32$. The interaction between lists and IRT type did not reach significance, $F(2, 57) = 2.71$, $MSE = 0.593$, $p = .08$. The IRT type effect was again significant or near-significant for each list [List 1, cluster M = 2.254 s , switch M = 2.609 s, $t(19) = 1.57$, $SD = 1.013$, $p = .067$ (one-tailed), $\eta^2 = .11$; List 2, cluster M = 1.996 s, switch M = 2.467 s, $t(19) = 2.32$, $SD = 0.487$, $p = .016$ (one-tailed), $\eta^2 = .22$; List 3, cluster M = 1.956 s, switch M = 3.056 s, $t(19) = 3.65$, $SD = 1.347$, $p = .0008$ (one-tailed), $\eta^2 = .41$].

Generally, IRTs become longer at later stages in people's recall sequences (Wixted & Rohrer, 1994). The present data also showed that trend. To examine the possible relation between output position and the shape cluster speed advantage, the normalized rank output position (ROP) within a trial for each IRT was calculated by the same formula as described for the previous FE study. Participants' mean cluster IRTs did not occur earlier within trials than their mean switch IRTs [mean normalized ROP of cluster IRTs = .4743, mean normalized ROP of switch IRTs = .0.4728, $t(32) = -0.12$, $SD = 0.076$, $p = .91$ (two-tailed)]. That was also true for the median normalized ROPs. Cluster IRTs were slightly later in output than switch IRTs for the items analysis, for both mean or median normalized ROPs, though again the difference was not significant. Therefore, the cluster speed advantage cannot be attributed to earlier output position. Indeed, as Figure 14 shows, cluster IRTs occurred later on the first trial.

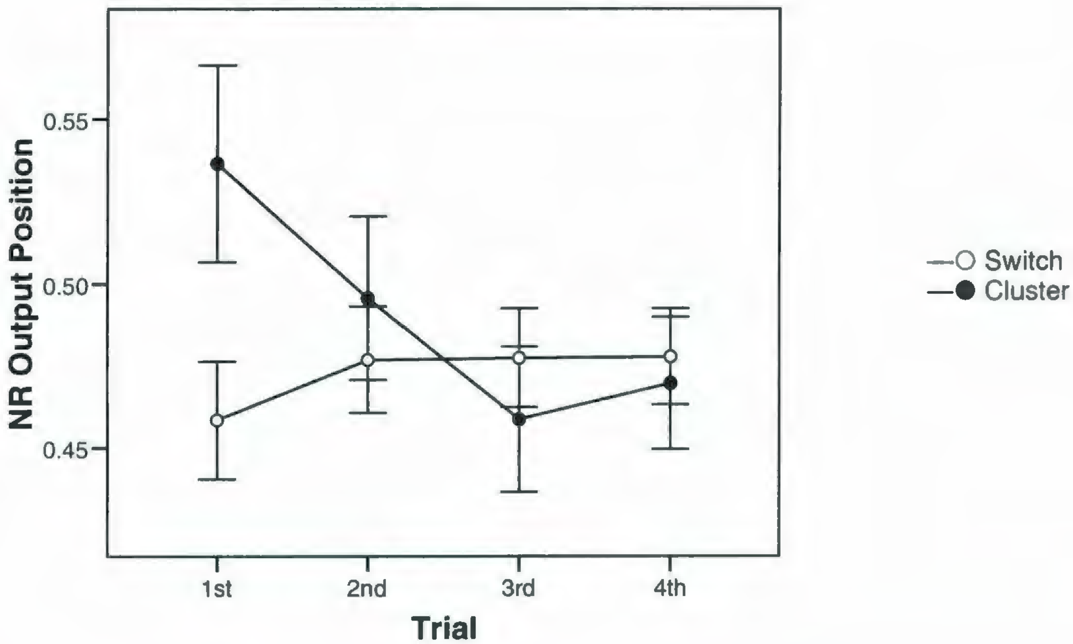

Figure 14. Normalized rank output position of cluster versus switch IRTs over trials (pooled data; $N = 1563$). Error bars represent ± 1 *SE.*

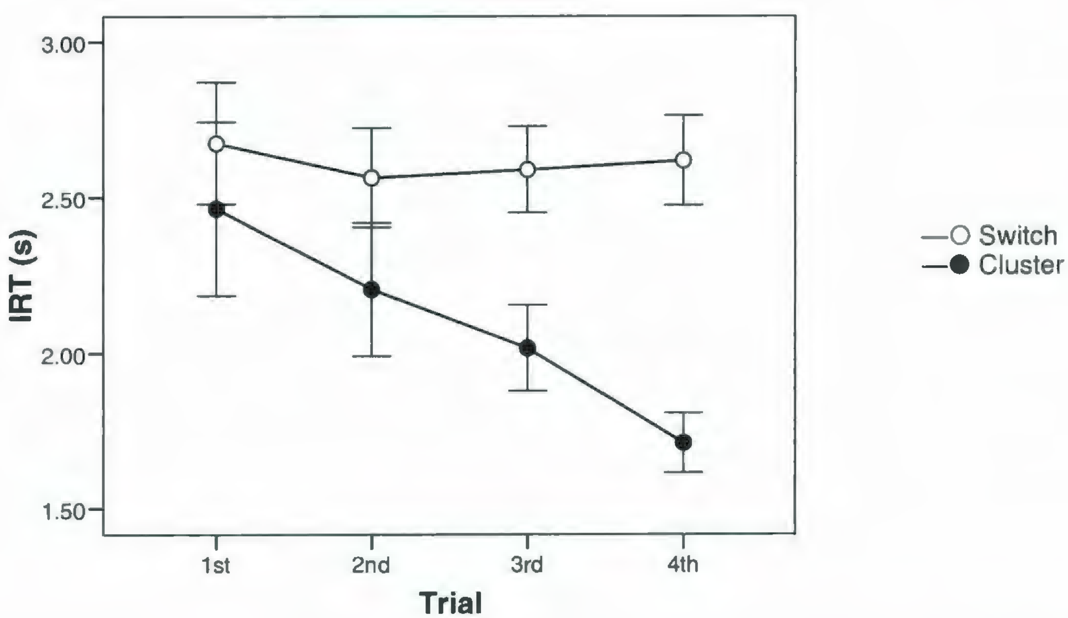

Figure 15. Cluster versus switch IRTs across trials (pooled data; $N = 1563$). Error bars represent ± 1 *SE.*

The shape cluster speed advantage for the first trial was obscured by high variability, as can be seen in Figure 15, but becomes clear on the later trials. There were noticeable primacy effects on the first trial, such that items from the early part of the study list tended to be recalled more, and tended to be recalled together in rapid succession, in the early phase of output. Perhaps because there were less-than-chance numbers of same-shape pairs in the study list, and perhaps because no same-shape pairings had been permitted to occur within the first three positions in the study list, people's tendency to recall the initial items together at the start of their recall sequence resulted in more switch IRTs earlier in the sequence, where IRTs tend to be faster. On the first trial, the shape clustering density (the number of cluster transitions divided by the total number of transitions) in the first two fifths of normalized rank output was much lower than in the later fifths of output. On subsequent trials, there were no obvious trends or systematic changes in clustering density across fifths of normalized rank output.

The speed advantage for clustering did not become clear until the middle to later stages of output (see Figure 16). The trend shown in Figure 16 occurred in all trials. On the fourth trial only (not shown), a clear cluster speed advantage also occurred in the first half of output, though again the advantage was much larger in the second half of output.

The cluster speed advantage was larger for longer shape clusters (runs), as shown in Figure 17. Note that a run length of 0 is a switch; a run length of 1 is a cluster containing two items of the same shape; a run length of 2 is a cluster containing three same-shape items; and so on.

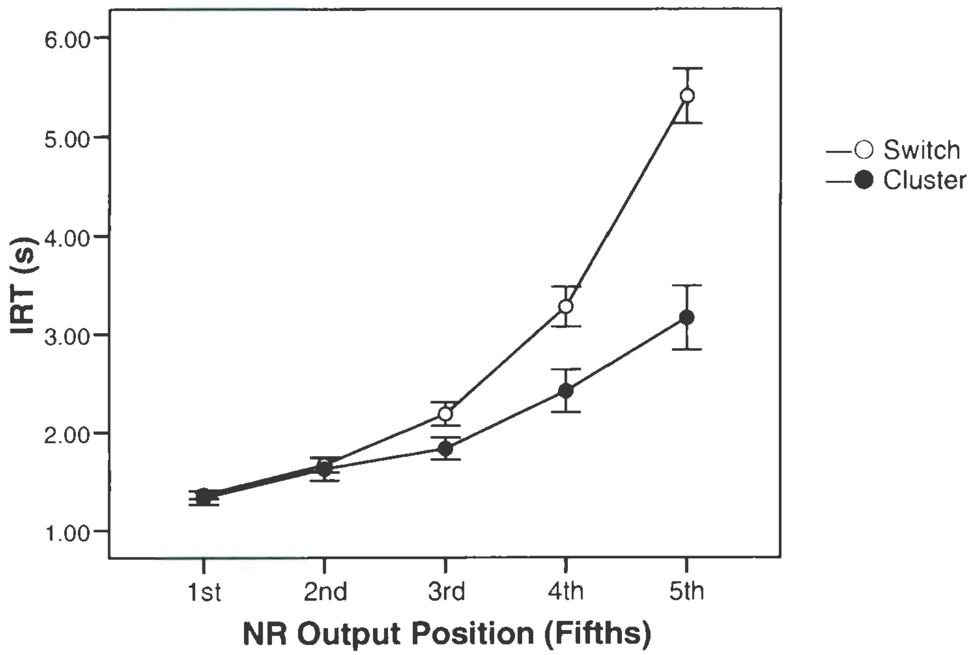

Figure 16. Cluster versus switch IRTs across successive fifths of normalized rank output position (pooled data; $N = 1563$). Error bars represent ± 1 *SE*.

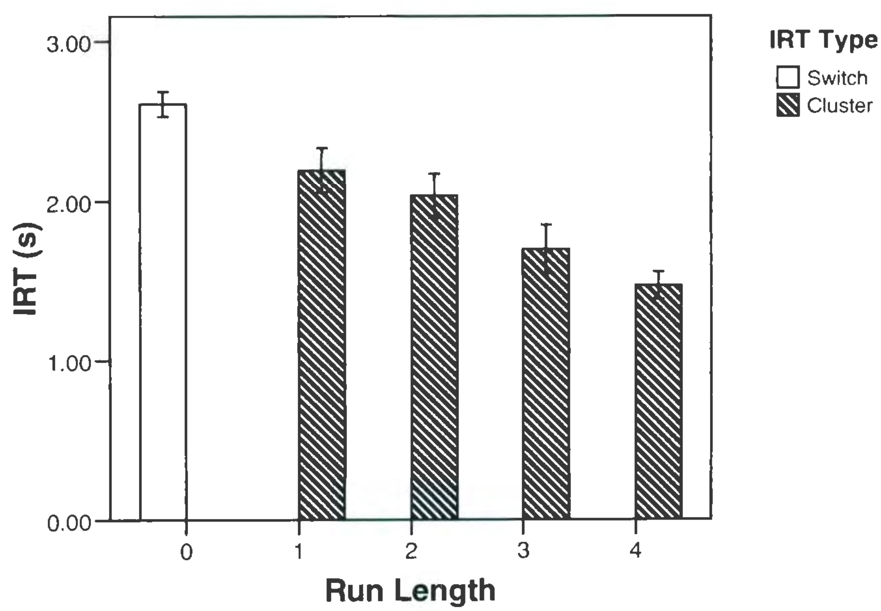

Figure 17. Mean IRTs (s) for shape cluster run lengths (pooled data; $N = 1563$). Note that run length is zero (0) for switches. Error bars represent ± 1 *SE*.

2.11.6. Post-Experiment Interview

The purpose of the post-experiment interview was to obtain qualitative information about the kinds of organizational schemes used by participants in the FR task. The interview was divided into three stages: (1) responding to open-ended questions about mental contents experienced while hearing and recalling the list, and a question about the reasons for improvement in recall over trials; (2) sorting and describing remembered groupings of the object names; and (3) responding to specific questions regarding shape similarity awareness. The classification and scoring schemes, and the results for each of the three stages, are shown below.

In their responses in Stages (1) and (2) of the interview, participants reported at least one type of memory content or scheme, which I will call a *mnemonic*. After I had obtained all of the post-task interview data, I classified these mnemonics according to the participants' labels and descriptions. I also made some basic distinctions (e.g., between semantic and episodic mnemonics) known to memory researchers. Table 24 provides the descriptions of the various kinds of mnemonics that participants reported. Inspection of the examples illustrates that the classifications are representative of the descriptions provided by participants.

Table 24

Descriptions of Mnemonics Reported in Post-Task Interview

Mnemonic	Description
E-Shape	*Experimenter-Defined Shape Category.* Reports a shape category (or categories) consistent with the experimenter-defined (E-defined) shape category, such as "cones" for *pylon, funnel, teepee.*
E-P Shape	*Experimenter-Participant-Defined Shape Category.* Reports a shape category or a grouping of items containing both E-defined and P-defined shapes. Can also include a mixture of two different E-defined shape categories that are deemed similar in shape (e.g., some long pointed cylinders grouped together with string-like things as "long").
P-Shape	*Participant-Defined Shape.* Reports the item(s) as being of a different shape than the E-defined shape (e.g., *log* construed as a squared timber or block of wood, grouped with *brick,* reportedly on the basis of this shape similarity; or *ribbon* construed in a loop shape, together with ring-shaped objects; or *bracelet* grouped with string-like things; etc.).
Unc-Shape	*Unclear-Items Grouped by Shape.* In the sorting task, the participant groups items together that happen to be of the same E-defined shape category, but, when asked, does not or can not give a description or label and/or does not know the basis of the grouping. The true basis for the grouping is unclear. The participant only indicates clearly that the items were grouped in memory.
Subst-pr	*Substitution—Primary.* An analogy is made between an object that is present in the testing room and a physically and/or semantically similar object named in the list. For example, the participant is reminded of *marker* by the pen in the room, or is reminded of *diamond ring* by his/her own ring, presently worn, which is not a diamond ring.
Subst-se	*Substitution—Secondary.* An analogy is made between an object that is not present in the testing room and a similar object named in the list (e.g., *pool cue* and *chopstick* are imaged as cross-pieces in a *teepee*; a *test tube* is likened to a shooter for drinking alcohol). The reported examples combined physical (e.g., shape) and semantic similarity.
OtherPerc	*Other Perceptual Similarity.* Reports any perceptual similarity, other than shape, between list objects (e.g., *pear* and *Life Saver* grouped together as having a sweet taste). This does not include Bodily

Perceptions, which is classified under Other (below).

Word-pr *Word-Primary Properties.* Mentions the surface properties of the word(s) such as phonology, spelling, or (rarely) word length, as distinctive characteristics or bases for grouping items (e.g., *frisbee* and *funnel* start with "f").

SpatialR *Spatial Relational.* Three types of spatial relations were reported. The most frequent were ad hoc or improvisational (e.g., otherwise unrelated objects placed next to one another, or balanced on top of one another, etc.), but were often consistent with shape- or structurally-based affordances (e.g., things put inside hollow objects, flat objects stacked, etc.). The second most frequent involved a list object (or objects) in some appropriate location, context, or scenario (e.g., *train car* on a track), but they did not occur in an episode or event (cf. Episodic). The least frequent involved imaginally laying out a row of separate abstract regions or "bins" for items in different shape categories, that could be surveyed from left-to-right or right-to-left in memorizing and in recalling the objects, shape category by shape category.

ActionR *Action Relational.* Two main types of action relations were reported. The first and most frequent were ad hoc or improvisational (e.g., throwing a *brick* and a *frisbee* at a *train car*), but sometimes were constrained by shape-based affordances (e.g., putting a *wreath* and a *hula hoop* around one's neck, crushing a *snowcone* down onto a *pylon*). The second type were more conventional and could include using a *marker* to write on a *playing card*, using a *quarter* in buying a *snowcone*, and so on. ActionR emphasized objects moving, rather than merely being positioned (cf. SpatialR).

Episodic *Episodic-Experiential.* Reports object(s) in the context of a specific autobiographical memory of a scenario or event (e.g., an event where friends or family members were smoking *cigars* and *playing cards*), or an imagined plausible scenario or event that is derived from one's own experience (e.g., putting a dog on a *leash* outside and seeing *worms* on the ground), or a specific individual object (e.g., remembers in detail a specific *licence plate* from an old car). Compared to Spatial and Action Relational, Episodic-Experiential schemes included more self-reference or reference to other people, were more grounded to specific experiences or situations, and had more reference to specific individual objects.

Semantic *Semantic Categorical or Relational.* Refers to a general semantic conceptual category or relation (e.g., *plate*, *snowcone*, and *juice box* grouped together as food-related; *casket* and *witch's hat* associated as spooky, Halloween-related things).

Personal	*Personal Knowledge.* Refers explicitly to self or some autobiographical fact in describing the basis for distinctiveness or organization of items. For examples, the participant plays *frisbee*, or knows a person named Mat (*mat*), or has a sister who wears a *ribbon* in her hair, or notes the "things I like"—*cigarette, juice box*. This mnemonic as reported was less contextual than Episodic and more self-referenced than Semantic.
Unc-Other	*Unclear-Items Grouped by Unknown Basis, or Unspecified Relation.* In describing the mnemonic, the participant says that items went together, were related, associated, linked in some way that is not specified or not known. In the context of the sorting task, this was the same as Unc-Shape, except that the items did not happen to be sorted according to the E-defined shape categories.
Imagery	In answering Questions 1, 2, or 3, the participant makes explicit use of words or phrasing such as "pictured," "visualized," "I saw," or "…was floating in front of me," and so on. Non-visual modalities were reported occasionally, but these were included under other categories. (No participant grouped items as imaged versus not imaged).
SerialOrg	*Serial Organization.* Participant reports having studied, recalled, or organized items according to their serial order of presentation. Participants who identified the approximate stage when this scheme occurred typically indicated having abandoned its use after the first trial.
VerbRept	*Verbal Repetition.* Participant reports having repeated a word over to themself in thought. Although this mnemonic may involve repetitive perceptuomotor imagery of the surface properties of the words, verbal repetition was not scored here as imagery.
NotPrvRecl	*Not Previously Recalled Items.* Participant reports having given special attention, priority, or importance to items that had not been recalled on previous trials, or indicates having kept track of what items had not yet been recalled either within or between trials. This mnemonic involves classifying items, during the encoding or retrieval phases, according to their status within the context of accomplishing the task goal.
PrvRecl	*Previously Recalled Items.* Similar to NotPrvRecl, except with reference to items already recalled.
UnsureNo	*Unsure of Memory Scheme, or No Definite Memory Scheme.* In answering Questions 1, 2, or 3, participant indicates being unsure or unaware of the memory scheme(s) used, or else states that no memory scheme was used.

Other	*Other Miscellaneous Memory Schemes.* Additional specific mnemonics were reported, but were not listed by more than three participants, and many were mentioned by only one participant. The specific Other mnemonics are listed here according to the part of the interview in which they occurred. *Hearing the list.* Other: Counting Items; Objects in Testing Room; Items Considered Strange or Unusual; Retracing Previous Words; Action Similarity; Item Linked to a Story Title or Subject. *Recalling the list.* Other: Bodily Perceptions; Counting Items; Relaxed or Passive. *Improvement in recall.* Other: Counting Items; Increased Concentration; Striving Toward Goal to Recall More; Relaxed or Passive. *Labelling the Sorted Units.* Other: Recalled Well or Not Well, Items Considered Strange or Unusual, Current Circumstances, Similar Material, Bodily Perceptions, and *Objects in the Testing Room. *(For details, see in Appendix I the Note under Table 13 for List 3).

2.11.7. Relative Frequencies of Reported Mnemonics

Two methods were used in scoring the frequencies of the mnemonics, one by number of participants, and the other by number of units. In the *number-of-participants* method, each mnemonic, when reported, was scored as a single report per question, even if the participant mentioned it more than once for that question. Thus, each mnemonic could be scored as having been reported by a maximum of 33 persons, per question. The contents of a sorted unit were scored more than once if the participant labelled or described them with more than one mnemonic. Table 25 displays the number of participants reporting each mnemonic for each of the open-ended questions and the sorting task.

The *number-of units* method dealt only with scoring the frequency and size of sorting task units. In this method, every unit in the entire pool of sorted units obtained from the 33 participants was counted and classified (*N* = 251 units). Again, contents of a sorted unit were scored more than once if the participant assigned more than one mnemonic label or description to them. Table 26 displays the distribution of the various mnemonic units of different sizes. Note that a unit of Size 1 is a single object name; Size 2 is a pairing of object names; Size 3 is a triplet of object names; and so on. A Size 1 unit was scored when the participant clearly and deliberately separated an object name from the others and attributed to it a mnemonic. Mnemonics associated with Size 1 units tended to refer to some distinguishing aspect of the item, or its association with subject matter outside of the list (e.g., other objects, contexts, etc.).

The level of agreement between the contents of participants' sorted units of Size 2 or larger and the contents of their item groupings in recall output was measured. For an example of scoring, if a sorted unit of Size 4 contained items (*a*, *b*, *c*, *d*), and only *c* and *a* occurred directly together in that person's recall output on at least one of the four trials, then 2 out of 4 items (50%) in that group showed agreement with the recall output. Items from each sorted unit were scored in this fashion, and the percentage of agreement (total sum of grouped items showing agreement, divided by the total sum of items grouped) was calculated for each person. The mean percentage of items within a person's sorted groupings that were directly conjoined at least once in the person's recall output sequences was 82 (*N* = 33; *SD* = 20.9; *Mdn* = 89), indicating adequate agreement.

In the sorting task, a total of 26 participants (79%) grouped at least a pair of items according to some kind of shape similarity (E-, E-P-, or P-defined shape). Table 25

shows that 24 of the participants (73%) grouped some items according to E-defined shapes. Among those 24 were 12 participants who also grouped by E-P-shape. Two of the 4 participants who grouped by P-shape did not group any other items by E- or E-P-shape. One third (11/33) of participants sorted at least one unit according to E-defined shape similarities without declaring a specific basis for the grouping (Unc-Shape).
In response to at least one of the post-task questions, 24 participants (73%) reported explicitly that their thinking or memory involved imagery (i.e., visualizing, picturing, etc.), with most of these reports given in response to the Hearing and Recalling questions, respectively (Table 25). An additional 8 participants (23%), who had not reported imagery explicitly, reported contents that implied imagery, including attention to object shape similarity, spatial relations, action relations, and/or episodic-experiential scenarios. Only one participant did not mention or imply imagery. That participant had focussed almost exclusively on organizing words according to common first letter/phoneme.

Those persons who did not report imagery ($n = 9$) did not show significantly above-chance mean clustering, nor shape cluster IRT speed advantage, but did show significantly above-chance mean SO-ARC scores. Those who reported imagery explicitly ($n = 24$) showed significantly above-chance mean clustering, a significant shape cluster IRT speed advantage, and significantly above-chance mean SO-ARC scores.

Most of the reported organizational schemes involved imagery. Explicit imagery terminology was reduced to only two reports when participants answered how they were able to improve their recall over trials. Participants' most common answer for the Improvement question was focussing on Items Not Previously Recalled. This mnemonic was not reported in people's labelling of their sorted groupings, perhaps because the set

Table 25

Numbers of Participants Reporting Each Mnemonic

Mnemonic	Hearing[a]	Recalling[b]	Improve.[c]	SortU[d]
E-Shape	7	5	4	24
E-P-Shape	4	3	3	12
P-Shape				4
Unc-Shape				11
Subst-pr				2
Subst-se	2			4
OtherPerc			1	3
Word-pr	4	4	4	6
SpatialR	10	6	3	10
ActionR	2	2	1	10
ExEpisod	4	4	2	5
Semantic	3	1	1	11
Personal	2	2	1	3
Unc-Other	2	1	1	17
Imagery	20	14	2	
SerialO	8	6		2
VerbRept	7			
PrvRecl	1	1	1	
NPrvRecl	7	8	9	
UnsureNo			6	
Other	6	6	5	11

Notes. Maximum *n* per cell = 33 for all mnemonic types, except for *Other*.
[a]Hearing the List, [b]Recalling the List, [c]Improvement over Trials, [d]Sorted Unit.

of items that were not previously recalled was transitory, changing within and between trials. The mnemonics reported more frequently in people's card-sorting (e.g., E-Shape, E-P-Shape, Semantic, SpatialR, and ActionR, etc.) were perhaps used in conjunction with the Items Not Previously Recalled mnemonic.

Table 25 shows that approximately one third of participants sorted at least one unit by SpatialR (10/33), ActionR (10/33), or Semantic (11/33), respectively. Slightly over half (17/33) of the participants sorted at least one unit according to an undeclared or unknown association that did not involve the E-shape similarities (Unc-Other). The most common explicitly reported perceptual mnemonics were shape similarities (excluding Unc-Shape); Spatial Relations; and Action Relations. Twenty-nine out of 33 participants (88%) reported at least one of those mnemonics explicitly.

Table 26 shows the distribution of all the reported units from the sorting task. Generally, there were more and larger units for E-defined shape categories (i.e., with a maximum size of 5). Some of the largest units were Spatial Relational, whereas Action Relational tended to occur in smaller units. The numbers and sizes of Semantic units were limited, probably due to the reduction of nuisance associations in the list. Serial Organization units were relatively scarce and small, probably due partly to the fact that items were presented in a different order on each trial. There were forty-four Size 1 units.

The participants' mean number of items sorted into units of Size 2 or larger was 14.5 (*SD* = 4.18; *Mdn* = 15; range 4 to 20). The mean number of sorted units of Size 2 or larger, per participant, was 6.3 (*SD* = 2.21; *Mdn* = 6; range 2 to 15; note that units were counted more than once if their contents were labeled by more than one mnemonic). The participants' mean number of sorted units were not significantly correlated with their mean recall, mean ARC, or mean IRT. Participants' mean sorted unit size, excluding units of Size 1, was 3.03 (*SD* = 0.841; *Mdn* = 2.83; range 2.0 to 5.0). Mean sorted unit size was similar to Number of Items Grouped in its pattern of correlations with those above-mentioned variables, though it was not as strongly correlated with those above-mentioned variables as was N Items Grouped (the latter variable is included in Table 27).

Table 26

Pooled Sums of Mnemonic Units by Type and Size

	Unit Size						
Unit Label	1	2	3	4	5	6+	Sum
E-Shape		26	9	10	15		60
E-P-Shape		2	3	3	2	2	12
P-Shape	2	4					6
Unc-Shape		15	2				17
Subst-pr	3	1					4
Subst-se	4	2	1				7
OthrPerc		3					3
Word-pr		5	4	1	1		11
SpatialR		6	2	1	3	5	17
ActionR	5	14	2				21
ExEpisod	8	5					13
Semantic	3	10	4	1	1		19
Personal	1	1			1		3
Unc-Othe	4	13	13	3	2		35
SerialO		4					4
Other	14	3	1		1		19
Sum	44	114	41	19	26	7	251

Notes. In the Size 6+ class, unit sizes ranged from 6 to 8. A Size 1 unit is a single item.

2.11.8. Shape Similarity Awareness

Participants' reported awareness of the E-defined shape similarities between objects in the list was scored as follows: 0 = no awareness reported, 1 = partial awareness reported, and 2 = "full" awareness reported (i.e., of all or most of the shape similarities). The "no awareness" participants (n = 9) did not report E-defined shape similarity awareness during any stage of the post-task interview before the first shape awareness question (Question 1) was asked. They did not label any grouping of cards explicitly according to the E- or E-P-defined shape similarities. In response to the shape awareness Question 1 or 3, they answered "no" or else mentioned that they were aware of P-defined shape similarities. The "partial awareness" participants (n = 18) mentioned at least some E-defined shape similarities, by which, during the post-task interview, they sorted from 2 to 10 items (M = 5.3). The "full awareness" participants (n = 6) reported an awareness of all or most of the shape similarities, by which they sorted from 13 to 20 items (M = 16.3). Many of the partial and full awareness participants also reported and sorted according to some E-P-defined shape similarities. Generally, among the participants who were able to estimate when they became aware of the shape similarities, the partial awareness participants recollected that the awareness arose on Trials 3 or 4, whereas the full awareness participants reported that the awareness arose on Trials 1 or 2. All of the 4 participants who mentioned the E-shape categories in response to the "Improvement of recall" question were full shape awareness participants. Of the three participants who mentioned E-P-shape categories in response to the "Improvement" question, two were partial awareness and one was a full awareness participant.

2.11.9. Shape Similarity Awareness, Organization, Recall, and IRTs

As Table 27 shows, higher E-shape similarity awareness was accompanied by faster mean IRTs, higher mean recall, higher numbers of items grouped in the post-task sort, and higher mean ARC. Shape similarity awareness levels did not differ across lists, $F(2, 30) = 1.35$, $MSE = 0.450$, $p = .27$, or number of same-shape pairs in the study list, $F(3, 29) = 2.32$, $MSE = 0.409$, $p = .10$. Note that the weak relation between number of

Table 27

Correlations Between Mean IRT, Mean ARC, and Other Variables

Variable	1	2	3	4	5	6
1. *M* IRT	----	**-.35** *.047*	**-.44** *.011*	**-.36** *.038*	**-.35** *.043*	**.04** *.843*
2. *M* Recall		----	**.44** *.011*	**.27** *.131*	**.39** *.026*	**.09** *.611*
3. N Items Grp.[a]			----	**.55** *.001*	**.63** *.00008*	**.20** *.274*
4. *M* ARC				----	**.62** *.0001*	**-.10** *.569*
5. Shape Aware.[b]					----	**.03** *.851*
6. *M* SO-ARC						----

Notes. Pearson correlations are in **bold**. Two-tailed *p*-values are in *italics*. All *N*s = 33.
[a]Number of Items Grouped in the post-experiment sorting task. [b]Shape Similarity Awareness.

same-shape pairs in the study list and shape awareness was such that there was less awareness when there were more same- pairs in the study list. The SO-ARC correlations show that neither shape similarity awareness levels nor ARC scores were associated with the carrying over of specific pairings of items from one recall trial to the next.

The frequencies of sorted unit labels laid out in Table 28 suggest that participants in the full awareness group showed less variety of mnemonics, as compared to the partial and no awareness groups. The full awareness group reported a mean of 2.8 *different* mnemonics in the sorting task, whereas the corresponding means for the partial and no awareness groups were 4.4 and 4.2, respectively.

The information in Table 28 confirms that participants who reported no awareness of the E-shape similarities showed low ARC scores. The relative numbers of persons sorting some items into Spatial and Action Relational units are about the same across levels of E-shape awareness. Not shown is the fact that participants in the two lower shape awareness groups produced larger and more numerous Spatial Relational groupings.

Compared to the other groups, the no awareness group had less mean items grouped in the sort, smaller mean sorted unit sizes, and the lowest percentage of participants reporting imagery (i.e., 56%, versus 72% and 100% reporting imagery in the partial and full awareness groups, respectively). Their mean level of recall was not less than that of the partial awareness group (see Table 28). The no-awareness group had reportedly focussed on other connections between the items, particularly on Semantic and Episodic connections, but also Spatial and Action Relations, and Word Properties. Five participants (56%) in the no awareness group reported either some P-defined shape similarities or substitutions involving object shape.

Seven participants (78%) in the no awareness group happened to sort a limited number of items (generally only 2 or 3 items) into E-defined shape categories, but the basis for these groupings was not declared (i.e., they were Unclear-Shape). Only one of the no awareness participants had no sorts by, or no reported awareness of, some type of shape similarity or shape-based substitution. That participant had focussed on organizing recall by the first letter/sound of the words (as noted in section 2.11.7).

The no awareness group showed slightly above-chance mean clustering and shorter cluster versus switch IRTs, but these results were not significant, with one-tailed

Table 28

Level of Shape Similarity Awareness, Performance, and Mnemonics

	Shape Similarity Awareness			
Measure	No	Partial	Full	
Participants (*n*)	9	18	6	
M Recall	12.8	12.6	16.0	
M N Items Grouped	11.4	14.3	19.5	
M Unit Size[a]	2.5	2.9	4.1	
M ARC	.06	.11	.53	
Unit Label				Sum
E-Shape	0	18	6	24
E-P-Shape	0	8	4	12
Unc-Shape	7	4	0	11
[b]PSha/Subst/OthP.	5	8	0	13
Word-pr	3	3	0	6
SpatialR	2	6	2	10
ActionR	3	4	3	10
ExEpisod	4	1	0	5
Semantic	5	5	1	11
Personal	2	0	1	3
Unc-Other	5	12	0	17
SerialO	1	1	0	2
Other	1	10	0	11
Sum	38	80	17	135

Notes. Each individual Unit Label was scored a maximum of once per participant, except for "Other" and the summed frequencies noted for PSha/Subst/OthP.
[a]Size 1 units were not included in mean Unit Size scores. [b]Summed frequencies for P-Shape, Primary and Secondary Substitutions, and Other Perceptual.

p-values of .18 and .25, respectively. The partial awareness group showed above-chance clustering [*M* ARC = .11; $t(17) = 3.41$, *SD* = 0.138, $p = .0017$ (one-tailed), $\eta^2 = .41$] and a shape cluster speed advantage [mean savings = 0.410 s; $t(17) = 1.92$, *SD* = 0.905, $p = .036$ (one-tailed), $\eta^2 = .18$]. The full awareness group showed strong clustering [*M* ARC = .53, $t(5) = 6.90$, *SD* = 0.189, $p = .0005$ (one-tailed), $\eta^2 = .90$], reaching a mean ARC of .82 on the fourth trial, and showed a strong mean shape cluster speed advantage [mean savings = 0.939 s, $t(5) = 3.50$, *SD* = 0.658, $p = .009$ (one-tailed), $\eta^2 = .71$].

2.11.10. Word Associations, Typicality, and Word Length

This analysis evaluates the possible contribution of shape word associations, shape typicality, and word length to the recall, clustering, and IRT results for items. Word association norms (Nelson, McEvoy, & Schreiber, 1998) were used to determine the shape word association strength for each item for which such information was available ($n = 33$ of 60). Association strength was the percentage of participants in the normative sample who gave a particular response word to a cue word. To find the association strength for a word's shape associates, the sum of all relevant shape-related names was obtained for each word. For example, for the target word *globe*, the sum of strengths for appropriate, shape-consistent associates such as "round," "ball," etc., was obtained. For 16 out of 33 words, the shape word association strength sum was zero (0). The mean sum shape word associative strength for the 33 words was 3.8% (*SD* = 5.11%; *Mdn* = 1.3%). A correlation analysis was conducted on these 33 items, for beginning- and end-word scores where appropriate. Using a two-tailed .05 criterion, shape word association strength was *not* significantly correlated with the following: Mean IRT, mean cluster IRT, mean switch IRT, mean switch – mean cluster IRT, number of times the item was recalled, number of times the item was clustered, cluster percentage (number of times the item was clustered / number of times the item was recalled), mean normalized rank output position, word length (number of letters in the word). Thus, shape word association strength was not a significant factor in explaining the recall, clustering, and IRT results.

Regarding nuisance associations, the Nelson et al. (1998) norms showed that "game" was given frequently in response to *volleyball, marble, card,* and *chess* (the latter was the closest available proxy for *chessboard*). Actually, these items clustered less often than other items, on average, within their respective shape categories. In List 3, "black" was listed in the norms as an associate for both *cannon* (1.4%) and *marker* (2%), though these items were not clustered with each other in recall. The List 1 items *marble* and *tile* were associated strongly to "floor" in the norms, but were also not clustered with each other.

As an item's shape word association strength sum increases, so does its mean object shape goodness (or typicality) rating [$r(31) = .40$, $p = .02$ (two-tailed)], though shape goodness was not significantly correlated with recall, clustering, or IRT duration. These typicality and word association results should be viewed with the caveat that there was a restriction of range for both variables (low levels of word association strength, and moderately high mean typicality).

Word length (number of letters in the word) was not correlated with most of the variables, but yielded an inconsistent pattern of associations with some performance variables. For the item data set ($N = 60$), for the ending-word analysis, word length was significantly or nearly significantly correlated with mean IRTs [$r(58) = .31$, $p = .02$ (two-tailed)], recall [$r(58) = .24$, $p = .06$ (two-tailed)], and cluster percentage [$r(58) = -.30$, $p = .02$ (two-tailed)]. However, for the beginning-word analysis, those relations were not significant between word length and the beginning-word mean IRTs [$r(58) = -.01$, $p = .94$ (two-tailed)], recall [$r(58) = .20$, $p = .12$ (two-tailed)], or cluster percentage [$r(58) = -.001$, $p = .996$ (two-tailed)]. For the pooled data ($N = 1563$), for either the end-word or the beginning-word analysis, there was no correlation between word length and IRT, clustering (as measured by IRT Type, cluster = 1 versus switch = 0), or recall (as measured by the number of IRTs per trial).

To examine IRT effects for long and short words, the distribution of items was split at the median word length ($Mdn = 6$; range 2 to 12), with words having six or less letters treated as short ($M = 4.9$; $n = 34$) and those with seven or more letters treated as long ($M = 8.5$; $n = 26$). Cluster IRTs were significantly briefer than switch IRTs for long words [beginning-word $t(25) = 2.19$, $SD = 1.048$, $p = .02$ (one-tailed), $\eta^2 = .16$; end-word $t(25) = 1.88$, $SD = 1.191$, $p = .036$ (one-tailed), $\eta^2 = .12$] and for short words [beginning-word $t(33) = 4.48$, $SD = 0.914$, $p = .00004$ (one-tailed), $\eta^2 = .38$; end-word $t(33) = 4.42$, $SD = 1.054$, $p = .00005$ (one-tailed), $\eta^2 = .37$]. Thus, while the effect is stronger for short words, it does occur for long words as well.

2.12. Free Recall Discussion

Consistent with predictions derived from the perceptual theories (Barsalou, 1999; Hebb, 1949), participants' recall outputs yielded significantly above-chance mean clustering of objects by shape category, along with faster mean shape cluster versus switch IRTs. Moreover, shape cluster IRTs were faster within longer shape cluster runs.

Participants' reported mnemonics and sorted groupings of the object names revealed the use of shape similarities, spatial and action relations, and other organizational schemes. The clustering, IRT, and reported mnemonic results are consistent with those obtained in the previous FR pilot study (*n* = 6, Mattless, 2003, unpublished).

Clustering and the cluster speed advantage were significant, or near-significant, within each of the three object name lists. This suggests that the shape clustering and IRT effects generalize across multiple sets of objects that have simple three-dimensional shapes. Lists did not differ significantly in clustering or IRT effects, though more items were recalled from List 2 as compared to List 1. Due to the small numbers of participants in each list group, one cannot rule out the possibility that there could be some differences between lists for clustering and IRT effects.

The number of same-shape pairs (0, 1, 2, or 3) in the study list was not significantly related to clustering in recall or IRT effects. Therefore, the clustering and IRT effects are not attributable to the below-chance numbers of same-shape pairs in the study list. Puff (1966), who used FR lists containing semantic categories, also found that including the chance level of same-category pairs did not produce higher clustering than a completely unblocked presentation.

Clustering was significant on all trials, including on the first trial for those participants who had adequate recall. Although the increase in ARC and MRR scores over trials was not significant, the percentage of participants who clustered above chance increased substantially from the first (51.5% above chance ARC) to second trial (73% above chance ARC). For the assessment of clustering, and given the number of categories in the list, the levels of recall in this study were more adequate than in those conditions in

Zimmer's (1989) study where shape clustering was usually near or below chance (see Appendix H). However, when Zimmer's levels of recall were clearly above adequate, as in the standard FR condition in Experiment 1 (standard instructions, Trial 2, *M* ARC = .13), and in Experiment 2 (standard instructions, *M* ARC = .12) his shape clustering results were only slightly lower than those reported here for List 2 (*M* ARC = .16) and List 3 (*M* ARC = .14). When mean recall levels are clearly above the minimum adequate level, significant mean clustering by E-defined shape categories does occur.[21]

Retrieval outputs showed normal, significantly above-chance levels of subjective organization, or inter-trial repetitions, as measured by the SO-ARC. Notably, SO-ARC scores were not positively correlated with ARC scores; that is, increases in E-shape category clustering did not involve increases in carrying over specific item pairings from one trial to the next.

On the first trial, participants generally showed primacy effects, recalling items from the early part of the study list together in a rapid burst in the early part of their recall output. In the first trial output, there was a lower cluster density in the first two fifths, as compared to the later three fifths, of output. On subsequent trials, cluster density did not show major changes across successive fifths of output. These latter results differ from those obtained by Bousfield and Cohen (1953, 1955) for FR of semantic categories, which showed a decline in cluster density in the later output positions. The reasons for this difference are unclear because of the many methodological differences between their studies and the present one. Despite those differences, their participants did show a markedly lower cluster density in the early phase, relative to the middle phase, of output in the single trial presentation condition, which is analogous to the low cluster density obtained in the early phase of the first trial in the present study.

Inspection of the pooled data (N = 1563) showed that the cluster IRT speed advantage was smaller on the first trial, particularly in the early stage of the first trial. This finding is consistent with that of the pilot FR study involving shape categories (Mattless, 2003, unpublished), and at least one other study involving recall of items from strong semantic categories (Thompson, 1978). In the present study, the primacy effect, plus the later occurrence of the shape IRTs relative to the shape switch IRTs, probably reduced the shape cluster speed advantage on the first trial. On subsequent trials, the cluster IRTs were not substantially earlier or later in output than switch IRTs, and the cluster speed advantage increased. Over all trials, the switching cost was non-existent or small in the early phase of output, and was progressively larger in the later stages of output. This trend is consistent with that observed in participants' FR of items from semantic categories (e.g., see the control condition in Patterson et al., 1971; and see Pollio et al., 1969) and subjectively organized units (Puff, 1972).

To address Engelkamp's and Zimmer's (1994) claim that there are amodal representations for some objects having strong shape word association (e.g., *ball – round*), and to address various other concerns about association strength, the object nouns selected for and used in this experiment had zero or very low levels of shape word association with the object name. As the above analyses showed, for the items for which shape word association information was available, there was no significant association between shape associate strength and clustering, IRT measures, or recall. However, one must interpret these results with caution because of the problems with word association as a construct (for discussion, see McRae & Boisvert, 1998; see especially p. 569). Also note that a finding of a word association effect would not contradict perceptual theories.

In the post-experiment interview, almost all participants reported explicitly or implied that their mnemonics involved imagery of objects and related information. The fact that most people reported imagery is consistent with other past research involving the recall of concrete nouns (e.g., see Richardson, 1999), but it is contrary to Engelkamp's and Zimmer's (1994) claim that perceptual representations are not activated in people's FR of object nouns under standard (non-imagery) instructions.

Participants' reported mnemonics and sorted groupings of object names mainly involved shape similarities, spatial and action relations, semantic relations and similarities, episodic scenarios, primary perceptual properties of the words, and alternative perceptual similarities. The validity of their reports is supported by the mean 82% agreement (*Mdn* = 89%) between their sorted groupings and their recall output groupings. In addition, increases in the number of items grouped in the sort and E-shape similarity awareness level, respectively, were significantly correlated with performance (shorter *M* IRT, increased recall, and increased ARC). Using an open-ended format for the questions (Hearing, Recalling, or Improvement) and the sorting task ensured that participants' responses reflected their own mental contents and organizing tendencies.

The finding that some semantic associations were reported is not unusual, despite the experimenter's efforts to reduce those nuisance associations in the lists here. Again, it is not possible to remove all nuisance associations. Semantic associations are still found by participants who recall items or sort items that are "unrelated" (e.g., see Schwartz & Humphreys, 1973).

The precise nature of the relation between E-shape similarity awareness and clustering is not clear from these data. The relationship may unfold during the task according to at least two possible scenarios. In one scenario, some participants, initially, may search consciously for any relations or similarities between the items, and then discover some shape similarities by which they, thereafter, organize their recall of the items. In another scenario, some participants may, initially, organize some items unconsciously by shape similarity, and then awareness of the shape similarities increases thereafter, leading to consciously guided organization by shape. A combination of those scenarios is also possible. In any case, shape similarity awareness could involve a heightened and more prolonged activation of the respective shape representations. Conscious awareness of the shape similarities would, presumably, allow participants the opportunity to manipulate the grouping of the items in working memory during encoding and retrieval, probably leading to the formation of longer clusters.

Interestingly, 9 of the 33 participants (27%) were reportedly not aware of any E-defined shape similarities. Some of these participants sorted some of the items according to episodic and semantic mnemonics, P-defined shape similarities and shape substitutions, spatial and action relations, or unintentionally or incidentally by E-defined shape, or by other mnemonics. These 9 participants showed E-shape clustering that was on average slightly, but not significantly, above chance; and their IRTs did not show a significant shape cluster speed advantage. The reported mnemonic and sorting results are consistent with those of the pilot FR study (Mattless, 2003, unpublished), where 2 of the 6 pilot participants remained unaware of the E-defined shape similarities until the post-experiment shape awareness questions were asked. Possibly, when participants engage in a more eclectic mix of mnemonics, and do not process any one type of information extensively, then they may be more likely to develop only partial or no awareness of the E-defined perceptual categories. In contrast to this outcome for single-feature perceptual categories, if participants applied an eclectic mix of mnemonics to the recall of lists containing concrete semantic categories, they could still show moderate significant clustering for the E-defined semantic categories because the items within those categories usually have more similarities and relations by which people can organize recall.

The finding that some participants do not become aware of the E-defined categories does have precedents. Frost's (1971) FR participants reportedly indicated having no awareness of the orientation categories. Loeb and DeNike (1969) found that 12% to 15% of their FR participants reported no awareness of the E-defined semantic categories (e.g., clothes, vehicles, animals, fruits, etc.) in the lists. Consistent with the present results, Loeb and DeNike (1969) also found that participants who reported using the E-defined categories to learn the list had higher clustering than those who reported mixed mnemonics (E-defined and non-E-defined categorical), and the mixed mnemonic participants in turn had higher clustering than those who reported no use of the E-defined categories.

Consistent with the usual practice in scoring clustering, in the present FR study only E-defined shape clusters were counted in the ARC and MRR score calculations. The instances of E-P- and P-defined shape groupings indicate that there was more shape similarity organization than was captured by the ARC and MRR measures. Reports of shape similarity-based analogies, such as *pool cue* imaged as a supporting pole in a *teepee*, obviously, could not be captured in the measurement of clustering. Such analogies are consistent with the "seeing as" and "imaging as" phenomena presented in the Introduction, where people image one object as another that is similar in shape. Overall, most participants (88%) explicitly reported at least one of the following perceptual or action mnemonics: shape similarity, spatial relations, and action relations. Moreover, 79% of all participants reported some kind of shape similarity mnemonic. These findings highlight the limitation of researchers' usual practice of reporting only the mean clustering, and of scoring clustering according to a single E-defined set of categories or relations. These findings now raise the possibility that participants who showed low mean E-defined perceptual clustering scores in previous research used some alternative perceptual mnemonics that were overlooked.

One limitation of the open-ended format used in the post-task interview is that participants may overlook, neglect to mention, or not remember some of the mnemonics they used during FR. For example, only one participant reported the mnemonic "Striving Toward the Goal to Recall More," even though such a goal was probably pursued by all participants. Likewise, the mnemonics involved in monitoring or keeping track of which items have or have not been recalled previously (PrvRecl and NPrvRecl) may be integral to FR task performance, though participants may not always report those mnemonics. In future research, the open-ended question format used here could be supplemented with a subsequently-administered mnemonic checklist (e.g., see Camp, Markley, & Kramer, 1983). A checklist could present all of the usual mnemonics, which would be equally available to be checked, or not checked, and ranked in importance by the participant. More FR studies, using a wider variety of items, and collecting mnemonic information in response to open-ended questions, are needed before a standard checklist can be constructed and justified.

In sum, the shape clustering, IRT, and reported mnemonic results confirm the predictions of perceptual theories. The mnemonic reports revealed additional perceptual similarity and relational schemes that were perhaps overlooked in previous research.

2.13. General Discussion

The present FE and FR studies provide the first clear evidence of secondary shape categorical organization in people's knowledge and LTM of everyday objects. On average, shape categorical clustering was significantly above chance, and shape cluster were faster than shape switch IRTs, in FE and FR retrieval. In their post-experiment reports, most participants mentioned that they used mental imagery and various perceptual schemes in their strategies and mnemonics. The shape clustering, IRTs, and post-experiment task reports confirmed the predictions derived from the perceptual theories (Barsalou, 1999; Hebb, 1949) but disconfirmed the claims of the amodal theories (Engelkamp & Zimmer, 1994; Shelton & Caramazza, 2001). That these effects occurred in the absence of any physical shape information suggests that the underlying representations and organization are based in people's knowledge or LTM systems. Upon retrieving relevant objects from LTM, people presumably activate object and shape representations, and these activations then temporarily facilitate the activation of other object representations having similar, overlapping shape representations (for further discussion of the hypothesized processes by which this occurs, see chapter 3).

In contrast to the present results, Zimmer's (1989) FR experiments did not show clearly above-chance clustering of nouns by the shape of the referent objects. Engelkamp and Zimmer (1994) cited Zimmer's (1989) findings to support their conclusion that memory for objects is not normally organized according to perceptual factors such as shape. The findings of the present FR study, which are based on adequate levels of recall, challenge Engelkamp's and Zimmer's (1994) claims. Engelkamp and Zimmer (1994) also claimed that memory for objects is not normally organized according to spatial relations or action relations. In the present FR study, most participants (88%) reported that they grouped items according to at least one of either shape similarity, spatial relations, or action relations. Most of these mnemonics involved grouping of items that were unrelated semantically. Semantically-unrelated items have now been shown to be organized according to shape category (in the present FR study and the pilot; also see Moar, 1977), depicted orientation (Frost, 1971), spatial relations (McNamara et al., 1989; Moar, 1977; Nida & Lange, 1997; Plumert, 1994; Plumert & Strahan, 1997; Stukuls, 1975), and action similarities (Koriat et al., 1998; Koriat & Pearlman-Avnion, 2003; cf. Zimmer & Engelkamp, 1989). Semantically-unrelated items are also clustered in FR or FE according to primary perceptuomotor similarities or relations, including similar or overlapping phonemes (Bousfield & Wicklund, 1969; Fagan, 1969; Troyer et al., 1998), similar abstract geometric figures (Bousfield et al., 1959; Jones-Gotman & Milner, 1977), similar melodies (Cutietta & Booth, 1996), spatial relations and proximities (Taylor & Tversky, 1992), and similar actions or gestures (Jason, 1985; Koriat et al., 1998; Koriat & Pearlman-Avnion, 2003; cf. Zimmer & Engelkamp, 1989).

The support of perceptual theories is not limited to evidence from laboratory tasks such as FE and FR. Consider again the "seeing as" and "imaging as" evidence presented in the Introduction, where people see or think of one object and then are reminded of another object that is similar in shape. The tendency for shape-similarity-based reminding would probably be established in the predominantly young adult participants long before they ever encountered the FE and FR tasks in this study. Shape-similarity-based reminding, analogies, and generalizations occur under a wide variety of natural and artificial conditions in adults, are evident in young children's normal learning of object names, and are even suggested in the behaviour of other species (Guthrie, 1993). Hence, it cannot be claimed that the effects of the same kind obtained here in FE and FR are only found in these experimental tasks. Rather, these shape clustering and IRT effects are just some examples from a long list of circumstances where such phenomena occur.

The experimental and observational evidence cited above challenges the claims of Engelkamp and Zimmer (1994), while supporting theories of memory that include perceptual and motor bases for inter-object similarities and relations (e.g., Barsalou, 1999, 2003; Hebb, 1949). Engelkamp and Zimmer (1994) proposed a multimodal theory of memory that was amodal with respect to categorical and relational representations, in which perceptual and motor information was normally not represented conceptually. The evidence reported and reviewed above suggests that human memory is a more extensively multimodal system that includes perceptual and motoric bases of organization.

2.13.1. Shape and Other Variables in Memory Organization

Although these studies focussed on object shape, the results also suggested that there were other major organizational tendencies in addition to, or confounded with, shape. Object shape is significantly confounded with semantic category in people's representations of everyday objects (also see Rosch et al., 1976). This confounding was reflected in the correlations between shape similarity and semantic transition in FE output and in the significant association between judges' shape and semantic classifications of objects. In FE output, "purely" shape cluster (SemSw—ShapeCl) and semantic cluster transitions (SemCl—ShapeSw), respectively, were each less frequent, and longer in duration, than semantic cluster-shape cluster transitions (SemCl—ShapeCl). The IRT results are consistent with reaction time results from the lexical decision and naming tasks of Flores d'Arcais et al. (1985) and Schreuder et al. (1984). The IRT results are also consistent with the reaction time results from the lexical and semantic decision studies of McRae and Boisvert (1998), where significant response time savings were obtained only when the semantically-related exemplars were also similar in other respects such as physical similarity. The classification and FE output results suggest that organization of objects in people's long-term knowledge involves a confounding between shape and semantic categories. Nonetheless, shape similarity effects are detected in FE output when the semantic variable is controlled statistically.

The classification study also revealed a significant confounding between semantic category and object size in people's long-term knowledge. There is a tendency for objects in the same semantic category to be of similar rated size. In FE output, within semantic clusters, inter-object transitions of low size difference were more frequent and faster than those of high size difference. In switching between semantic categories, however, there was no significant speed advantage for transitions of low size difference, suggesting that object size—at least for performance within the FE domains used here—is only a significant organizing factor within semantic categories of objects. On the other hand, Moar's (1977) FR results suggested that people do organize retrieval of semantically-unrelated objects according to size. Due to the confounds in Moar's (1977) presentation (e.g., *house* shown in a larger picture and *cup* shown in a smaller picture), it would be necessary to carry out another FR study, using only object names presented one at a time in random (or unblocked) sequence, to determine if people have a significant tendency to organize retrieval of unrelated objects by known size.

The classifications and FE outputs demonstrated consistent associations between the shape, size, and semantic variables, confirming the results of previous studies that had shown substantial overlap of object features within concrete semantic categories (McRae et al., 1997; McRae et al., 1999; Rosch & Mervis, 1975). When feature overlap between exemplars is reduced, semantic categorical clustering in recall is also reduced (e.g., see Cissè & Heth, 1989).

People's memory for everyday objects also appears to be organized strongly according to spatial location/context—the predominant strategy reported for FE here and for other FE tasks in other research (Bond et al., 1985; Vallée-Tourangeau et al., 1998; Walker & Kintsch, 1985; Williams & Hollan, 1981; Williams & Santos-Williams, 1980). Following from the present line of research, future studies using FE might require participants, in the post-task procedure, to classify the listed objects according to location. It would then be possible to examine (a) the amount of confounding between object location, semantic category, and other variables, in people's classifications, and (b) the amount of clustering and the influence on IRT durations attributable to location.

Whereas FE participants' reported use of spatial location/context strategies involved predominantly a search for objects in known locations from everyday life, FR participants' reported spatial relational mnemonics involved mostly ad hoc scenarios constructed for the purpose of combining unrelated objects into memorable groupings. In FE, participants seemed to recreate an experience similar to actually searching physical environments for objects. In FR, some participants at retrieval presumably attempted to reinstate the spatial groupings that they formed during study of the list.

2.13.2. Imagery and Visuospatial Processing

Most of the FE and FR participants used explicit imagery terminology in describing the strategies and mnemonics that they had used in retrieving object items, despite the fact that there were no instructions to use imagery. This finding, together with the finding that most of the participants reported perceptual schemes that implied the use of imagery, such as spatial location/contextual, shape similarity, and spatial and action relations, suggests that people activated perceptual and motor representations during the tasks. Past research has established that people activate appropriate perceptual and motor areas of the brain during imagery tasks (for reviews, see Barsalou, 2008; Kosslyn, 1994; Kosslyn, Ganis, & Thompson, 2001; Richardson, 1999). Imagery phenomena are reported to occur naturally, without imagery instructions, when people recall concrete nouns (Richardson, 1999), when they freely emit semantic and ad hoc category items (Rende, 1999, cited in Rende et al., 2002; Vallée-Tourangeau et al., 1998; Walker & Kintsch, 1985, Exp. 1), the names of former acquaintances (Williams & Hollan, 1981; Williams & Santos-Williams, 1980), or events that typically occur within common scripts (e.g., "going to a restaurant for a meal," etc.; Walker & Kintsch, 1985, Exp. 2). Imagery phenomena also occur when people verify object properties (Solomon & Barsalou, 2004; also see Kan, Barsalou, Solomon, Minor, & Thompson-Schill, 2003), and when they list object properties (Wu & Barsalou, 2003, see "neutral" condition).

The notion that concrete semantic FE performance involves visuospatial processing is also supported by results from studies using dual-task methodology. For example, performing a concurrent visuospatial task reduces production in a semantic fluency task (e.g., animals, fruits and vegetables, clothing, furniture) more than in a first-letter fluency task, and the reverse pattern of results is obtained for semantic and letter fluency tasks when the concurrent tasks are articulatory suppression (Rende et al., 2002) or sequential finger-tapping on a computer keyboard with the right hand by right-handed participants (Martin, Wiggs, Lalonde, & Mack, 1994). This suggests that at least some of the major concrete semantic FE tasks rely more on visuospatial processing, while first-letter FE tasks rely more on the specialized perceptuomotor systems involved in the surface or primary aspects of speech processing or sequencing of simple manual actions.

The widespread findings of imagery reported in FE and recall tasks suggest that Engelkamp and Zimmer (1994) were incorrect when they claimed that perceptual representations are not activated under standard non-imagery instruction conditions involving words as stimuli or responses. Note also that Pylyshyn's (1984, 2002) amodal "tacit knowledge" hypothesis purports to explain how people process and respond to imagery instructions, but the present research did not use imagery instructions.

2.13.3. Accessing Object Knowledge Activates Perceptual and Motor Areas

The present FE and FR results can be added to the growing body of evidence supporting perceptually-based theories of concrete knowledge representation. When people name objects, actions, and perceptual properties, even when only word stimuli are presented, there is significant activation in appropriate perceptual and motor cortical regions (for reviews, see Barsalou, 2008; Barsalou, Simmons, Barbey, & Wilson, 2003; Martin et al., 2000; Pulvermüller, 1999; and see Thompson-Schill, Oliver, Brainard, & Robison, 2003). Appropriate areas of perceptual cortex are activated when people access their knowledge of the sensory properties of objects in response to object names (for a review, see Martin, 2007; also see Goldberg, Perfetti, & Schneider, 2006a, 2006b; Simmons et al., 2007) and when they generate images of known entities from object names (Handy et al., 2004, see noun condition) or from animal names (Lambert, Sampaio, Scheiber, & Mauss, 2002).

Participants incur response time costs when they must switch between (as compared to within) stimulus property modalities in verifying the properties of noun referents (Pecher, Zeelenberg, & Barsalou, 2003; Simmons, Pecher, Hamann, Zeelenberg, & Barsalou, 2003). Property verification (Solomon & Barsalou, 2001) experiments show that response times are shorter when there is higher (versus lower) shape similarity between prime and target named objects. In sentence verification, in verifying whether a depicted object was mentioned as a noun in a previous sentence context, people show a response time cost if the depicted object appears in an exemplar shape that is different from what would be expected from the previous sentence context (Zwaan, Stanfield, & Yaxley, 2002).

Areas of the brain that process shape information during perception are also activated when people process object word or shape word stimuli, with no actual or depicted object or shape information present. For example, people show increased activation of some areas of the parietal cortex, including areas analogous to those found to be selective for three-dimensional shapes in primates (e.g., see Sakata et al., 1999), when they judge shape descriptions of object noun referents (e.g., "curved" versus "straight," in Oliver & Thompson-Schill, 2003). When people judge or rate the shape complexity of object noun referents, areas involved in perceiving object shape, including the inferior parietal cortex, and the lateral occipital complex (LOC), become more highly active (Newman, Klatzky, Lederman, & Just, 2005). When people passively read shape words such as "square," "arc," and so on, as compared to when they read colour words, some areas involved in perceiving objects and shape are activated, such as the middle temporal gyrus and the anterior fusiform gyrus, and there is also heightened activation in specific premotor, inferior prefrontal, and dorsolateral prefrontal areas (Pulvermüller & Hauk, 2006). When people listen passively to lists of spoken nouns referring to everyday objects, there is coherent activation between posterior inferior temporal, posterior parietal, and occipital sites (von Stein, Rappelsberger, Sarnthein, & Petsche, 1999)—that is, areas that are active in perceiving objects and shape. The evidence from these different studies suggests that if people in the present FR study had mentally rehearsed the shape category verbal labels, or explicitly attended to referent object shape, activation of shape-selective cortical areas would probably have occurred.

2.13.4. Abstraction and Categorical Representation

A common misconception about perceptual models of cognition is that they cannot, in principle, achieve categorical abstraction (Pylyshyn, 1973; also see Adams & Campbell, 1999; Gabora, 1999; Ohlsson, 1999). This misconception has been perpetuated, in part, by the repeated claims that perceptual representations are merely recordings of stimulus patterns. These claims are unsupported and, indeed, run contrary to the explicitly stated assumptions of modern perceptual theorists. Perceptual theorists assume that the representations are normally schematic, incomplete, and dynamic, not exact rigid copies of individual stimuli (e.g., Barsalou, 1999, 2003; Hebb, 1949, 1968). These properties are attributed to not only categorical representations, but also to relational and contextual representations. Some researchers have proposed that many aspects of knowledge may be organized not according amodal schemes, but according to abstract spatial and/or spatio-temporal schemas (e.g., Barsalou, 1999, pp. 590-594; Edelman & Breen, 1999; Hock & Schmelzkopf, 1980; Mandler, 1992). Goldstone and Barsalou (1998; also see Barsalou, 2003) have described numerous ways in which abstraction can be achieved in perceptual systems.

Primary clustering in FR and FE tasks, cited previously, provides physical trace evidence of the hypothesized abstraction or categorization effects that are perceptual and/or motoric. The shape clustering results of the present studies can be taken as evidence of perceptually-based abstraction and classification having occurred in the absence of the referent object stimuli. The evidence, presented in the Introduction, dealing with "seeing as" and shape similarity-based reminding, shows that perceptual abstraction and stimulus generalization occur in a broad range of circumstances (Guthrie, 1993), not just in experiments. All of this evidence suggests that categorical abstraction does occur in perceptually-based representational systems.

2.13.5. Assessing the Limits of Perceptual Organizational Effects

In the retrieval of objects and entities, significant mean clustering may not necessarily occur for all perceptual variables under all conditions. For that matter, neither does significant clustering occur for all semantic categories (e.g., see Shuell's 1969 review). Differences in levels of clustering probably have less to do with the perceptual versus semantic distinction and more to do with other factors. These factors may include the strength of within-category relations (e.g., spatial, action, etc.), amount of feature overlap between items, item dominance, contrast and alignment between categories, coherence within categories, lexical ambiguity, the number and strength of competing organizational schemes, and possibly other factors, such as the availability of various categories to conscious attention.

It would be unreasonable for any theorist to postulate that memory of objects is organized according to features that are not likely to be important in experience. The extent to which a person's knowledge of everyday objects is organized according to shape—or any other perceptual feature—may depend on how much experience the person has with the objects, the quantity and diversity of objects known to the person, the particular semantic domain of objects, and other factors. It may be the case, for example, that a person's knowledge of objects is organized more strongly according to shape if the person has extensive expertise and skill in interacting with, classifying, or using those objects. Due to their vocation or hobby, some people, such as visual artists, may attend to the shapes of things more systematically, more often, and more thoroughly than others. Historical changes and cultural differences in the kinds of objects that are available, and the modes by which people interact with them, could affect the extent of the influence of object shape on memory organization.

The generality of these FE and FR results is somewhat limited by the sampling of objects, which included mostly those of simple shape. To support broader generalizations about the organization of object memory, an *objects* FE task (e.g., like the practice FE task used here) could be used in future research, using analyses similar to those used in the present research, to determine the relative contribution of shape similarity, size difference, location, semantic transition, and other item content variables, to clustering and IRT effects. Likewise, a FR study could be conducted using longer lists of objects that were sampled at random from a large population of reasonably familiar objects. It would be possible to assess the respective levels of shape, semantic categorical, spatial relational, and other types of clustering and IRT effects, based on participants' post-task classifications. Whereas judges in the classification study sorted objects in the master lists according to specific experimenter-defined criteria (shape, size, or semantic categories), it would also be useful to have judges freely classify large numbers of object items, and then the extent to which the judges used each type of classification could be measured. These methods would allow for broader generalization about the organization of memory of objects.

Although some object shape clustering appears to occur for most participants and items when the recall level is above adequate, it does not necessarily occur for all participants, or for all items. Yet nearly all participants in the FE and FR experiments here reported at least one example of a perceptual strategy or mnemonic, besides imagery, for retrieving items. These perceptual schemes may arise whenever the items to be retrieved are concrete objects or entities.

Manipulating the FE or FR orienting instructions could also produce informative results. For example, preparing participants to list or study-and-recall the items with the intention of later drawing them, or manipulating them, might increase the object shape activations, particularly of the defining boundaries of the objects. Other conditions could examine the effect of instructions that ask participants to imaginally trace the boundaries of the objects.

Now that the present FE and FR studies are complete, it would first be prudent to replicate the results. Subsequently, additional studies could be carried out examining the role of shape similarity or difference in various paradigms designed to assess LTM and knowledge organization. For example, the effectiveness of object shape similarity and difference could be studied in FR distinctiveness (or von Restorff) paradigms, including the usual measures of item recall frequency and priority, but also the measurement of clustering, IRTs, and post-task reported mnemonics. Lexical and semantic decision experiments, using items generated from the present research, could also be carried out to determine whether or not there is a significant reaction time savings for similarly-shaped as compared to differently-shaped prime-target pairs of object noun referents.

2.14. Summary Conclusions

One can now make sense of the apparent discrepancy that arose due to claims that, while semantic clustering was widespread, perceptual clustering did not occur. This appeared to support the claims of the amodal theory of organization. As consideration of past evidence showed, however, the apparent discrepancy or discontinuity turned out to be chiefly a difference in the level of clustering. Perceptual clustering did occur when methodological problems were minimized. Perceptual clustering occurred at lower levels than for most of the concrete semantic categories that have been tested thus far, at least partly due to the increased amounts of naturally confounded similarities and relations between items within semantic categories, and the greatly reduced similarities and relations within perceptual categories as they are normally studied. Increased clustering associated with increased overlap between category members, whether according to perceptual or semantic similarity, can be explained parsimoniously within perceptual theories according to a common set of principles and mechanisms (as discussed in the next chapter). Perceptually-based memory organization is not an exception. Rather, it conforms to the general principles of similarity and relational organization. Perceptual similarity and perceptual relational organizational tendencies exist; they can be demonstrated in classical memory organization tasks; and the phenomena are not limited to experimental tasks, as shown by the shape analogy and reminding-by-shape-similarity phenomena that occur in a wide variety of circumstances.

In the next chapter, I attempt to develop a principled scientific explanation of clustering. I work toward an explanation of the hypothesized neural mechanisms underlying shape categorical clustering in retrieval.

References for Chapters 1 and 2

Adams, F., & Campbell, K. (1999). Modality and abstract concepts. *Behavioral and Brain Sciences, 22*, 610.

Ammons, A. R. (1974). *Sphere: The form of a motion.* New York: W. W. Norton & Co.

Anderson, J. R., & Bower., G. H. (1980). *Human associative memory: A brief edition.* Hillsdale, NJ: Lawrence Erlbaum.

Anderson, R. E., & Helstrup, T. (1993). Visual discovery in mind and on paper. *Memory & Cognition, 21*, 283-293.

Anglin, J. M. (1983). Extensional aspects of the preschool child's word concepts. In Th. B. Seiler, & W. Wannenmacher, (Eds.), *Concept development and the development of word meaning* (pp. 247-266). New York: Springer-Verlag.

Ashcraft, M., Kellas, G., & Needham, S. (1975). Rehearsal and retrieval processes in free recall of categorized lists. *Memory & Cognition, 3*, 506-512.

Baldwin, D. A. (1989). Priorities in children's expectations about object label reference: Form over color. *Child Development, 60*, 1291-1306

Baldwin, D. A. (1992). Clarifying the role of shape in children's taxonomic assumption. *Journal of Experimental Child Psychology, 54*, 392-416.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577-660.

Barsalou, L. W. (2003). Abstraction in perceptual symbol systems. *Philosophical Transactions of the Royal Society of London, Series B, 358*, 1177-1187.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*. 617-645. Preview retrieved December 21, 2007 from http://psychology.emory.edu/cognition/barsalou/papers/Barsalou_ARP_2008_grounded_cognition.pdf

Barsalou, L. W., Simmons, W. K., Barbey, A., & Wilson, C. D. (2003). Grounding knowledge in modality-specific systems. *Trends in Cognitive Sciences, 7*, 84-91.

Biederman, I. (1987). Recognition-by-components: A theory of human image interpretation. *Psychological Review, 94*, 115-147.

Biederman, I., & Cooper, E. E. (1991). Priming contour-deleted images: Evidence for intermediate representations in visual object recognition. *Cognitive Psychology, 23*, 393-419.

Biederman, I., & Ju, G. (1988). Surface versus edge-based determinants of visual recognition. *Cognitive Psychology, 20*, 38-64.

Bjorklund, D. F., (1988). Acquiring a mnemonic: Age and category knowledge effects. *Journal of Experimental Child Psychology, 45*, 71-87.

Bond, C. F., Jones, R. L., & Weintraub, D. L. (1985). On the unconstrained recall of acquaintances: A sampling-traversal model. *Journal of Personality and Social Psychology, 49*, 327-337.

Booth, G. D., & Cutietta, R. A. (1991). The applicability of verbal processing strategies to recall of familiar songs. *Journal of Research in Music Education, 39*, 121-131.

Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *The Journal of General Psychology, 49*, 229-240.

Bousfield, W. A., & Abramczyk, R. (1966). Sequential ordering in repeated free recall as a function of the length of the stimulus word list. *Canadian Journal of Psychology, 20*, 427-434.

Bousfield, W. A., Berkowitz, H., & Whitmarsh, G. (1959). Associative clustering in the recall of minimally meaningful geometric designs. *Canadian Journal of Psychology, 13*, 281-287.

Bousfield, W. A, & Cohen, B. H. (1953). The effects of reinforcement on the occurrence of clustering in the recall of randomly arranged associates. *The Journal of Psychology, 36*, 67-81.

Bousfield, W. A., & Cohen, B. H. (1955). The occurrence of clustering in the recall of randomly arranged words of different frequencies of usage. *Journal of General Psychology, 52*, 83-96.

Bousfield, W. A., Cohen, B. H., & Whitmarsh, G. (1958). Associative clustering in the recall of words of different taxonomic frequencies of occurrence. *Psychological Reports, 4,* 39-44.

Bousfield, W. A., & Puff, C. (1964). Clustering as a function of response dominance. *Journal of Experimental Psychology, 67,* 76-79.

Bousfield, W. A., & Sedgewick, C. H. W. (1944). An analysis of sequences of restricted associative responses. *The Journal of General Psychology, 30,* 149-165.

Bousfield, W. A., & Wicklund, D. A. (1969). Rhyme as a determinant of clustering. *Psychonomic Science,* 16, 183-184.

Bower, G. H., & Clark-Meyers, G. (1980). Memory for scripts with organized versus randomized presentation. *British Journal of Psychology, 71,* 369-377.

Braunstein, M. L., Hoffman, D. D., & Saidpour, A. (1989). Parts of visual objects: An experimental test of the minima rule. *Perception, 18,* 817-826.

Brennan, R. L., & Light, R. J. (1974). Measuring agreement when two observers classify people into categories not defined in advance. *British Journal of Mathematical and Statistical Psychology, 27,* 154-163.

Brewer, W. F. (1999). Perceptual symbols: The power and limitations of a theory of dynamic imagery and structured frames. *Behavioral and Brain Sciences, 22,* 611-612.

Brook, R. J., & Stirling, W. D. (1984). Agreement between observers when the categories are not specified in advance. *British Journal of Mathematical and Statistical Psychology, 37,* 271-282.

Camp, C. J., Markley, R. P, & Kramer, J. J. (1983). Naïve mnemonics: What the "do nothing" control group does. *American Journal of Psychology, 96,* 503-511.

Caramazza, A., Hersh, H., & Torgerson, W. S. (1976). Subjective structures and operations in semantic memory. *Journal of Verbal Learning and Verbal Behavior, 15,* 103-117.

Caramazza, A., Hillis, A. E., Rapp, B., & Romani, C. (1990). The multiple semantic hypothesis: Multiple confusions? *Cognitive Neuropsychology, 7,* 161-189.

Chan, A. S., Butters, N., Paulsen, J. S., Salmon, D. P., Swenson, M. R., & Maloney, L. T. (1993). An assessment of the semantic network in patients with Alzheimer's disease. *Journal of Cognitive Neuroscience, 5,* 254-261.

Cissè, D, & Heth, D. (1989). An evaluation of the differential encoding and feature overlap accounts of typicality effects in free recall. *Canadian Journal of Psychology, 43,* 359-368.

Cofer, C. N., Bruce, D. R., & Reicher, G. M. (1966). Clustering in free recall as a function of certain methodological variations. *Journal of Experimental Psychology, 71,* 858-866.

Cohen, B. H. (1963). Recall of categorized word lists. *Journal of Experimental Psychology, 66,* 227-234.

Connor, C. E. (2004). Shape dimensions and object primitives. In Leo M. Chalupa and John S. Werner (Eds.), *The visual neurosciences* (Vol. 2, pp. 1080-1089). Cambridge, Mass: MIT Press.

Cummins, C. (2006). Critical values for the Durbin-Watson test: 5% significance level. Retrieved November, 2006, from http://www.stanford.edu/~clint/bench/dwcrit.htm

Cutietta, R.A., & Booth, G.A. (1996). The influence of metre, mode, interval type and contour in repeated melodic free recall. *Psychology of Music, 24,* 222-236.

Dabady, M., Bell, M., & Kihlstrom, J. F. (1999). Person memory: Organization of behavior by traits. *Journal of Research in Personality,* 33, 369-377.

Damasio A. R. (1990). Category-related recognition defects as a clue to the neural substrate of knowledge. *Trends in Neurosciences, 13,* 95-99.

Deese, J. (1959). Influence of inter-item associative strength upon immediate free recall. *Psychological Reports, 5,* 305-313.

Diesendruck, G., & Bloom, P. (2003). How specific is the shape bias? *Child Development, 74,* 168-178.

Edelman, S., & Breen, E. M. (1999). On the virtues of going all the way. *Behavioral and Brain Sciences, 22,* 614.

Eimas, P. D., & Quinn, P. C. (1994). Studies on the formation of perceptually based basic-level categories in young infants. *Child Development, 65,* 903-917.

Einstein, G. O., & Hunt, R. R. (1980). Levels of processing and organization: Additive effects of individual-item and relational-processing. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 588-598.

Engelkamp, J., & Zimmer, H. D. (1994). *Human memory: A multimodal approach.* Toronto: Hogrefe and Huber.

Fagan, J. F., III. (1969). Clustering of related but nonassociated items in free recall. *Psychonomic Science, 16,* 92-93.

Farah, M. J. (1990). *Visual agnosia: Disorders of object recognition and what they tell us about normal vision.* Cambridge, MA: MIT Press.

Finke, R. A., Pinker, S., & Farah, M.J. (1989). Reinterpreting visual patterns in mental imagery. *Cognitive Science, 13,* 51-78.

Finke, R. A., & Slayton, K. (1988). Explorations of creative visual synthesis in mental imagery. *Memory & Cognition, 16,* 252-257.

Fiske, A. P. (1995). Social schemata for remembering people: Relationships and person attributes that affect clustering in free recall of acquaintances. *Journal of Quantitative Anthropology, 5,* 305-324.

Flores d'Arcais, G. B., Schreuder, R., & Glazenborg, G. (1985). Semantic activation during the recognition of referential words. *Psychological Research, 47,* 39-49.

Fodor, J. A. (1983). *Modularity of mind: An essay on faculty psychology.* Cambridge, MA: MIT Press.

Frei, A. (2003). Notes on temporal autocorrelation and regression. Retrieved November, 2006, from http://www.geography.hunter.cuny.edu/~afrei/gtech702_fall03_webpages/notes_temporal_autocorrelation.htm

Friendly, M. (1979). Methods for finding graphic representations of associative memory structures. In C. R. Puff, (Ed.), *Memory organization and structure.* (pp. 85-129). New York: Academic Press.

Frost, N. (1971). Clustering by visual shape in the free recall of pictorial stimuli. *Journal of Experimental Psychology, 88,* 409-413.

Frost, N. (1972). Encoding and retrieval in visual memory tasks. *Journal of Experimental Psychology, 95,* 317-326.

Fuster, J. (1995). *Memory in the cerebral cortex: An empirical approach to neural networks in the human and nonhuman primate.* Cambridge, MA: Bradford, MIT Press.

Gabora, L. (1999). Grounded in perceptions yet transformed into amodal symbols. *Brain and Behavioral Sciences, 22,* 630-631.

Gardner, H., Winner, E., Bechhofer, R., & Wolf, D. (1978). The development of figurative language. In K. E. Nelson (Ed.), *Children's language* (Vol. 1, pp. 1-38). New York: Gardner Press.

Gelman, S. A., Croft, W., Fu, P., Clausner, T., & Gottfried, G. (1998). Why is a pomegranate an *apple*? The role of shape, taxonomic relatedness, and prior lexical knowledge in children's overextension of *apple* and *dog*. *Journal of Child Language, 25,* 267-291.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition, 23,* 183-209.

Gershberg, F. B., & Shimamura, A. P. (1995). Impaired use of organizational strategies in free recall following frontal lobe damage. *Neuropsychologia, 13,* 1305-1333.

Goldberg, R. F., Perfetti, C. A., & Schneider, W. (2006a). Distinct and common cortical activations for multimodal semantic categories. *Cognitive, Affective, and Behavioral Neuroscience, 6*, 214-222.

Goldberg, R. F., Perfetti, C. A., & Schneider, W. (2006b). Perceptual knowledge retrieval activates sensory brain regions. *Journal of Neuroscience, 26*, 4917-4921.

Goldstone, R. L., & Barsalou, L. W. (1998). Reuniting perception and conception. *Cognition, 65*, 231-262.

Gollin, E. S., & Sharps, M. J. (1988). Facilitation of free recall by categorical blocking depends on stimulus type. *Memory & Cognition, 16*, 539-554.

Graesser, A. C., II., & Mandler, G. (1978). Limited processing capacity constrains the storage of unrelated sets of words and retrieval from natural categories. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 86-100.

Graham, S. A., Kibreath, C.S., & Welder, A. N. (2004). Thirteen-month-olds rely on shared labels and shape similarity for inductive inferences. *Child Development, 75*, 409-427.

Gregory, R. L. (1973). The confounded eye. In R. L. Gregory & E. H. Gombrich, (Eds.), *Illusion in art and nature* (pp. 49-95). New York: Charles Scribner's Sons.

Gruenewald, P., & Lockhead, G. (1980). The free recall of category examples. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 225-240.

Guthrie, S. E. (1993). *Faces in the clouds: A new theory of religion* New York: Oxford University Press.

Halpern, A. R. (1984). Organization in memory for familiar songs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 496-512.

Handy, T. C., Miller, M. B., Schott, B., Shroff, N. M., Janata, P., Van Horn, J. D., et al., (2004). Visual imagery and memory: Do retrieval strategies affect what the mind's eye sees? *European Journal of Cognitive Psychology, 16*, 631-652.

Hasselhorn, M. (1992). Task dependency and the role of category typicality and metamemory in the development of an organizational strategy. *Child Development, 63*, 202-214.

Hays, W. L. (1994). *Statistics* (5th ed.). Toronto: Harcourt Brace.

Hayward, W. G. (1998). Effects of outline shape in object recognition. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 427-440.

Hebb, D. O. (1949). *The organization of behavior.* New York: John Wiley & Sons.

Hebb, D. O. (1968). Concerning imagery. *Psychological Review, 75*, 466-477.

Helstrup, T., & Anderson, R. E. (1996). On the generality of mental construction in imagery: When bananas become smiles. *European Journal of Cognitive Psychology, 8*, 275-293.

Henley, N. M. (1969). A psychological study of the semantics of animal terms. *Journal of Verbal Learning and Verbal Behavior, 8*, 176-184.

Hirtle, S. C., & Jonides, J. (1985). Evidence of hierarchies in cognitive maps. *Memory & Cognition, 13*, 208-217.

Hock, H. S., & Schmelzkopf, K. F. (1980). The abstraction of schematic representations from photographs of real-world scenes. *Memory & Cognition, 8*, 543-554.

Hodge, M. H., & Otani, H. (1996). Beyond category sorting and pleasantness rating: Inducing relational and item-specific processing. *Memory & Cognition, 24*, 110-115.

Hoffman, D. D., & Richards, W. A. (1984). Parts of recognition. *Cognition, 18*, 65-96.

Hubert, L. (1977). Nominal scale response agreement as a generalized correlation. *British Journal of Mathematical and Statistical Psychology, 30*, 98-103.

Hudson, R. (1968). Category clustering as a function of level of information and number of stimulus presentations. *Journal of Verbal Learning and Verbal Behavior, 7*, 1106-1122.

Hudson, R. (1969). Category clustering for immediate and delayed recall as a function of recall cue information and response dominance variability. *Journal of Experimental Psychology, 82*, 575-577.

Humphreys, G. W., & Forde, E. M. E. (2001). Hierarchies, similarity, and interactivity in object recognition: "Category-specific" neuropsychological deficits. *Behavioral and Brain Sciences, 24*, 453-509.

Hunt, E., & Love, T. (1972). How good can memory be? In A. H. Melton & E. Martin, (Eds.), *Coding processes in human memory* (pp. 237-260). Washington, DC: V. H. Winston & Sons.

Hunt, R. R., & Einstein, G. O. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning and Verbal Behavior, 20*, 497-514.

Hunt, R. R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory and Language, 32*, 421-445.

Imai, M., Gentner, D., & Uchida, N. (1994). Children's theories of word meaning: The role of shape similarity in early acquisition. *Cognitive Development, 9*, 45-75.

Jahnke, J. (1965). Primacy and recency effects in serial position curves of immediate recall. *Journal of Experimental Psychology, 70*, 130-132.

Janson, H. W. (1973). Chance images. In Wiener, P. P., (Ed.), *Dictionary of the history of ideas: Studies of selected pivotal ideas: Vol. 1. Abstraction in the formation of concepts to design argument* (pp. 340-353). New York: Charles Scribner's Sons.

Jason, G. W. (1985). Gesture fluency after focal cortical lesions. *Neuropsychologia, 23*, 463-481.

Jenkins, J. J., & Russell, W. A. (1952). Associative clustering during recall. *The Journal of Abnormal and Social Psychology, 47*, 818-821.

Jones-Gotman, M., & Milner, B. (1977). Design fluency: The invention of nonsense drawings after focal cortical lesions. *Neuropsychologia, 15*, 653-673.

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition, 24*, 103-109.

Kahana, M. J., & Wingfield, A. (2000). A functional relation between learning and organization in free recall. *Psychonomic Bulletin & Review, 7*, 516-521.

Kan, I. P., Barsalou, L. W., Solomon, K. O., Minor, J. K., & Thompson-Schill, S. L. (2003). Role of mental imagery in a property verification task: fMRI evidence for perceptual representations of conceptual knowledge. *Cognitive Neuropsychology, 20*, 525-540.

Katz, A. N. (1981). Knowing about the sensory properties of objects. *Quarterly Journal of Experimental Psychology, 33A*, 39-49.

Katz, A. N. (1983). Dominance and typicality norms for properties: Convergent and discriminant validity. *Behavior Research Methods & Instrumentation, 15*, 29-38.

Kausler, D. H. (1974). *The psychology of verbal learning and memory*. New York: Academic Press.

Kay, D., & Anglin, J. (1982). Overextension and underextension in the child's expressive and receptive speech. *Journal of Child Language, 9*, 83-98.

Kellas, G., Ashcraft, M., Johnson, N., & Needham, S. (1973). Temporal aspects of storage and retrieval in free recall of categorized lists. *Journal of Verbal Learning and Verbal Behaviour, 12*, 499-511.

Khan, M, & Paivio, A. (1988). Memory for schematic and categorical information: A replication and extension of Rabinowitz and Mandler (1983). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 558-561.

Kobasigawa, A., & Orr, R. (1973). Free recall and retrieval speed of categorized items by kindergarten children. *Journal of Experimental Child Psychology, 15*, 187-192.

Kolb, B., & Whishaw, I. (1990). *Fundamentals of human neuropsychology* (3rd ed.). New York: W. H. Freeman.

Koriat, A., & Pearlman-Avnion, S. (2003). Memory organization of action events and its relationship to memory performance. *Journal of Experimental Psychology: General, 132*, 435-454.

Koriat, A., Pearlman-Avnion, S., & Ben-Zur, H. (1998). The subjective organization of input and output events in memory. *Psychological Research, 61*, 295-307.

Kosslyn, S. M. (1994). *Image and brain: The resolution of the imagery debate.* Cambridge, MA: Bradford, MIT Press.

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience, 2*, 635-642.

Kroes, W., & Libby, W. (1971). Relative power of taxonomic, semantic-differential, and sense impression categories for the organization of recall. *The Journal of General Psychology, 85*, 165-169.

Kroes, W., & Libby, W. (1973). Relative power of selected encoding categories for the organization of free recall in children. *Journal of Genetic Psychology, 122*, 9-15.

Lambert, S., Sampaio, E., Scheiber, C., & Mauss, Y. (2002). Neural substrates of animal mental imagery: Calcarine sulcus and dorsal pathway involvement—An fMRI study. *Brain Research, 924*, 176-183.

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development, 3*, 299-321.

Lederman, S. J., & Klatzky, R. L. (1990). Haptic classification of common objects: Knowledge-driven exploration. *Cognitive Psychology, 22*, 421-459.

Loeb, J. W., & DeNike, L. D. (1969). Prediction of organization in free recall from reported learning strategy. *Psychological Reports, 25*, 631-636.

Mandler, G., & Dean, P. J. (1969). Seriation: Development of serial order in free recall. *Journal of Experimental Psychology, 81*, 207-215.

Mandler, J. M. (1992). How to build a baby: II. Conceptual primitives. *Psychological Review, 99*, 587-604.

Martin, A. (2001). Functional neuroimaging of semantic memory. In R. Cabeza, & A. Kingstone, (Eds.), *Handbook of functional neuroimaging of cognition* (pp. 153-186). Cambridge, MA: Bradford, MIT Press.

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology, 58,* 25-45.

Martin, A., Ungerleider, L. G., & Haxby, J. V. (2000). Category specificity and the brain: The sensory/motor model of semantic representations of objects. In Gazzaniga, M. S. (ed.), *The new cognitive neurosciences* (2nd ed., pp. 1023-1037). Cambridge, MA: Bradford, MIT Press.

Martin, A., Wiggs, C. L., Lalonde, F., & Mack, C. (1994). Word retrieval to letter and semantic cues: A double dissociation in normal subjects using interference tasks. *Neuropsychologia, 32,* 1487-1494.

McDaniel, M. A., Einstein, G. O., & Lollis, T. (1988). Qualitative and quantitative considerations in encoding difficulty effects. *Memory & Cognition, 16,* 8-14.

McNamara, T. P., Hardy, J. K., & Hirtle, S. C. (1989). Subjective hierarchies in spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 211-227.

McRae, K., & Boisvert, S. (1998). Automatic semantic similarity priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 558-572.

McRae, K., Cree, G. S., Westmacott, R., & de Sa, V. R. (1999). Further evidence for feature correlations in semantic memory. *Canadian Journal of Experimental Psychology, 53,* 360-373.

McRae, K., de Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General, 126,* 99-130.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100,* 254-278.

Melkman, R., & Deutsch, H. (1977). Memory function as related to developmental changes in bases of organization. *Journal of Experimental Child Psychology, 23*, 84-97.

Melkman, R., Tversky, B., & Baratz, D. (1981). Developmental trends in the use of perceptual and conceptual attributes in grouping, clustering, and retrieval. *Journal of Experimental Child Psychology, 31*, 470-486.

Michela, J. L. (1990). Within-person correlational design and analysis. In C. Hendrick and M. S. Clark, (Eds.), *Research methods in personality and social psychology: Vol. 11. Review of personality and social psychology* (pp. 279-311). Newbury Park, CA: Sage Publications.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81-97.

Miotto, E. C., Savage, C. R., Evans, J. J., Wilson, B. A., Martins, M. G. M., Iaki, S., et al. (2006). Bilateral activation of the prefrontal cortex after strategic semantic cognitive training. *Human Brain Mapping 27*, 288-295.

Moar, I. (1977). Spatial organization in the free recall of pictorial stimuli. *Quarterly Journal of Experimental Psychology, 29 A*, 699-708.

Murphy, M. D. (1979). Measurement of category clustering in free recall. In C. R. Puff, (Ed.), *Memory organization and structure* (pp. 51-83). New York: Academic Press.

Murphy, M. D., & Puff, C. R. (1982). Free recall: Basic methodology and analyses. In C. R. Puff, (Ed.), *Handbook of research methods in human memory and cognition* (pp. 99-128). New York: Academic Press.

Nairne, J. S., Riegler, G. L., & Serra, M. (1991). Dissociative effects of generation on item and order retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 702-709.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. Retrieved December 21, 2007, from http://w3.usf.edu/FreeAssociation/

Neter, J., Kutner, M. H., Nachtsheim, C. J., & Wasserman, W. (1996). *Applied linear statistical models* (4th ed.). Toronto: Irwin.

Newman, S. D., Klatzky, R. L., Lederman, S. J., & Just, M. A. (2005). Imagining material versus geometric properties of objects: an fMRI study. *Cognitive Brain Research, 23*, 235-246.

Nida, R. & Lange, G. (1997). Young children's use of spatial cues to organize and recall object names. *Journal of Genetic Psychology, 158*, 457-465.

Ohlsson, S. (1999). Selecting is not abstracting. *Behavioral and Brain Sciences, 22*, 617.

Oliver, R. T., and Thompson-Schill, S. L (2003). Dorsal stream activation during retrieval of object size and shape. *Cognitive, Affective, & Behavioral Neuroscience, 3*, 309-322

Palmer, S. E. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA: MIT Press.

Patterson, K. E., Meltzer, R. H., & Mandler, G. (1971). Inter-response times in categorized free recall. *Journal of Verbal Learning and Verbal Behavior, 10*, 417-426.

Pecher, D., Zeelenberg, R., & Barsalou, L. (2003). Verifying different-modality properties for concepts produces switching costs. *Psychological Science, 14*, 119-124.

Pecher, D., Zeelenberg, R., & Raaijmakers, J. G. W. (1998). Does pizza prime coin? Perceptual priming in lexical decision and pronunciation. *Journal of Memory and Language, 38*, 401-418.

Pellegrino, J. W. (1971). A general measure of organization in free recall for variable unit size and internal sequential consistency. *Behavior Research Methods & Instrumentation, 3*, 241-246.

Perlmutter, M., & Ricks, M. (1979). Recall in preschool children. *Journal of Experimental Child Psychology, 27*, 423-436.

Perrig, W. J., & Hofer, D. (1989). Sensory and conceptual representations in memory: Motor images that cannot be imaged. *Psychological Research, 51,* 201-207.

Peterson, M. A. (1993). The ambiguity of mental images: Insights regarding the structure of shape memory and its function in creativity. In B. Roskos-Ewoldsen, M. J. Intons-Peterson, and R. E. Anderson, (Eds.), *Imagery, creativity, and discovery: A cognitive perspective* (pp. 151-185). New York: North-Holland.

Peterson, M. A. (2003). Overlapping partial configurations in object memory: An alternative solution to classic problems in perception and recognition. In M. A. Peterson and G. Rhodes, (Eds.), *Perception of faces, objects, and scenes: Analytic and holistic processes* (pp. 269-294). New York: Oxford University Press.

Peterson, M. A., & Rhodes, G., (Eds.). (2003). *Perception of faces, objects, and scenes: Analytic and holistic processes.* New York: Oxford University Press.

Plumert, J. M. (1994). Flexibility in children's use of spatial and categorical organizational strategies in recall. *Developmental Psychology, 30,* 738-747.

Plumert, J. M., & Strahan, D. (1997). Relations between task structure and developmental changes in children's use of spatial clustering strategies. *British Journal of Developmental Psychology, 15,* 495-514.

Pollio, H. R. (1964). Composition of associative clusters. *Journal of Experimental Psychology, 67,* 199-208.

Pollio, H. R., Kasschau, R. A., & DeNise, H. E. (1968). Associative structure and the temporal characteristics of free recall. *Journal of Experimental Psychology, 76,* 190-197.

Pollio, H. R., Richards, S., & Lucas, R. (1969). Temporal properties of category recall. *Journal of Verbal Learning and Verbal Behavior, 8,* 529-536.

Poulin-Dubois, D., Frank, I., & Graham, S. A., & Elkin, A. (1999). The role of shape similarity in toddlers' lexical extension. *British Journal of Developmental Psychology, 17,* 21-36.

Puff, C. R. (1966). Clustering as a function of the sequential organization of the stimulus list words. *Journal of Verbal Learning and Verbal Behavior, 5,* 503-506.

Puff, C. R. (1972). Temporal properties of organization in recall of unrelated words. *Journal of Experimental Psychology, 92,* 225-231.

Puff, C. R. (Ed.). (1979). *Memory organization and structure.* New York: Academic Press.

Puff, C. R., Murphy, M. D., & Ferrera, R. A. (1977). Further evidence about the role of clustering in free recall. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 742-753.

Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences, 22,* 253-336.

Pulvermüller, F. & Hauk, O. (2006). Category-specific conceptual processing of color and form in left fronto-temporal cortex. *Cerebral Cortex, 16,* 1193-1201.

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin, 80,* 1-24.

Pylyshyn, Z. W. (1984). *Computation and cognition: Towards a foundation for cognitive science.* Cambridge, MA: Bradford, MIT Press.

Pylyshyn, Z. W. (2002). Mental imagery: In search of a theory. *Behavioral and Brain Sciences, 25,* 157-238.

Quinn, P. C., Eimas, P. D., & Tarr, M. J. (2001). Perceptual categorization of cat and dog silhouettes by 3- to 4-month-old infants. *Journal of Experimental Child Psychology, 79,* 78-94.

Rabinowitz, M., & Mandler, J. M. (1983). Organization and information retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 430-439.

Rakison, D., & Butterworth, G. (1998). Infants' use of object parts in early categorization. *Developmental Psychology, 34,* 49-62.

Rende, B., Ramsberger, G., & Miyake, A. (2002). Commonalities and differences in the working memory components underlying letter and category fluency tasks: A dual-task investigation. *Neuropsychology, 16, 309-321.*

Richardson, J. T. E. (1999). *Imagery.* Hove, England: Psychology Press/Taylor & Francis.

Robertson, C., & Ellis, H. D. (1987). Estimating the effects of various clustering schemes on recall order. *British Journal of Mathematical and Statistical Psychology, 40,* 1-19.

Roenker, D. L., Thompson, C. P., & Brown, S. C. (1971). Comparison of measures for the estimation of clustering in free recall. *Psychological Bulletin, 76,* 45-48.

Rosch, E. H. (1973). Natural categories. *Cognitive Psychology, 4,* 328-350.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7,* 573-605.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8,* 382-439.

Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General, 126,* 211-227.

Rubin, D. C., & Olson, M. J. (1980). Recall of semantic domains. *Memory & Cognition, 8,* 354-366.

Sakata, H., Taira, M., Kusunoki, M., Murata, A., Tanaka, Y., & Tsutsui, K. (1999). Neural coding of 3D features of objects for hand action in the parietal cortex of the monkey. In G. W. Humphreys, J. Duncan, & A. Treisman, (Eds.), *Attention, space, and action: Studies of cognitive neuroscience.* New York: Oxford University Press.

Samuelson, L. K., & Smith, L. B. (2000). Children's attention to rigid and deformable shape in naming and non-naming tasks. *Child Development, 71,* 1555-1570.

Schwartz, R. M., & Humphreys, M. S. (1973). Similarity judgments and free recall of unrelated words. *Journal of Experimental Psychology, 101*, 10-15.

Shelton, J. R., & Caramazza, A. (2001). The organization of semantic memory. In B. Rapp, (Ed.), *The handbook of cognitive neuropsychology: What deficits reveal about the human mind* (pp. 423-444). Philadelphia, PA: Psychology Press.

Shepard, R. N. (1957). Multidimensional scaling of concepts based on restricted sequences of responses. *American Psychologist, 12*, 440-441. [Abstract]

Shepard, R. N., & Cermak, G. W. (1973). Perceptual-cognitive explorations of a toroidal set of free-form stimuli. *Cognitive Psychology, 4*, 351-377.

Shepard, R. N., & Chipman, S. (1970). Second-order isomorphism of internal representation: Shapes of states. *Cognitive Psychology, 1*, 1-17.

Schreuder, R., Flores d'Arcais, G. B., & Glazenborg, G. (1984). Effects of perceptual and conceptual similarity in semantic priming. *Psychological Research, 45*, 334-354.

Shuell, T. J. (1969). Clustering and organization in free recall. *Psychological Bulletin, 72*, 353-374.

Simmons, K., Pecher, D., Hamann, S. B., Zeelenberg, R., & Barsalou, L. W. (2003). fMRI evidence for modality-specific processing of conceptual knowledge in six modalities. Meeting of the Society for Cognitive Neuroscience, New York, March 2003. Retrieved December 21, 2007, from http://www.psychology.emory.edu/cognition/barsalou/papers/Simmons_et_al_CNS_poster_2003_multimodal_simulation.pdf

Simmons, K. W., Ramjee, V., Beauchamp, M. S., McRae, K., Martin, A., & Barsalou, L. W. (2007). A common neural substrate for perceiving and knowing about color. *Neuropsychologia, 45*, 2802-2810.

Sjolander, K., & Beskow, J. (2004). *Wavesurfer* [Computer Software]. Stockholm, Sweden: Centre for Speech and Technology, at KTH. Retrieved December, 2004, from http://www.speech.kth.se/wavesurfer/

Small, S. L., Hart, J., Nguyen, T., & Gordon, B. (1995). Distributed representations of semantic knowledge in the brain. *Brain, 118*, 441-453.

Smith, L. B. (1987). Commentary: Perceptual relations and perceptual language. *Monographs of the Society for Research in Child Development, 52,* (Serial No. 215, pp. 94-102). Lawrence E. Marks, Robin J. Hummeal, Marc H. Bornstein, authors of this volume, "Perceiving Similarity and Comprehending Metaphor." Series Editors: Robert N. Emde, Marshall M. Haith, Nancy Rumbaugh Whitesell.

Sodian, B. D., Schneider, W., & Perlmutter, M. (1986). Recall, clustering, and metamemory in young children. *Journal of Experimental Child Psychology, 41,* 395-410.

Solomon, K. O., & Barsalou, L. W. (2001). Representing properties locally. *Cognitive Psychology, 43,* 129-169.

Solomon, K. O., & Barsalou, L. W. (2004). Perceptual simulation in property verification. *Memory & Cognition,* 32, 244-259.

Stine, E. L., Benham, A. E., & Smith, A. D. (1987). Visual and semantic organization in picture recall. *Bulletin of the Psychonomic Society, 25,* 89-91.

Stukuls, H. (1975). Clustering as determined by exposure time and spatial arrangement of objects. *Psychological Reports, 37,* 159-166.

Taylor, H. A., & Tversky, B. (1992). Descriptions and depictions of environments. *Memory & Cognition, 20,* 483-496.

Thomas, C. B., Ross, D. C., & Freed, E. S. (1964). *An index of Rorschach responses. Studies on the psychological characteristics of medical students –I.* Baltimore, MD: The Johns Hopkins Press.

Thompson, C. P. (1978). Evidence for learning-to-cluster as a retrieval strategy. *American Journal of Psychology, 9,* 115-124.

Thompson-Schill, S. L., Oliver, R. T., Brainard, D. H., & Robison, S. G. (2003). Retrieval of color information activates color perception areas. Poster presentation. Retrieved December 21, 2007, from http://www.psych.upenn.edu/stslab/newcolorposter.ppt.pdf

Torrance, E. P. (1974). Directions manual and scoring guide: Figural test booklet B. In E. P. Torrance, *Torrance tests of creative thinking* (Revised). Benseville, Ill.: Scholastic Testing Service.

Tröster, A. I., Warmflash, V., Osorio, I., Paolo, A. M., Alexander, L. J., & Barr, W. B. (1995). The roles of semantic networks and search efficiency in verbal fluency performance in intractable temporal lobe epilepsy. *Epilepsy Research, 21*, 19-26.

Troyer, A. K., Moscovitch, M., Winocur, G., Alexander, M. P., & Stuss, D. (1998). Clustering and switching on verbal fluency: The effects of focal frontal- and temporal-lobe lesions. *Neuropsychologia, 36*, 499-504.

Tulving, E. (1962). Subjective organization in free recall of "unrelated" words. *Psychological Review, 69*, 344-354.

Tulving, E., & Donaldson, W. (Eds.). (1972). *Organization of memory*. New York: Academic Press.

Tversky, B. (1989). Parts, partonomies, and taxonomies. *Developmental Psychology, 25*, 983-995.

Tversky, B., & Hemenway, K. (1984). Objects, parts, and categories. *Journal of Experimental Psychology: General, 113*, 169-193.

Underwood, B. J., & Richardson, J. (1956). Some verbal materials for the study of concept formation. *Psychological Bulletin, 53*, 84-95.

Ungerleider, L. G. (1995). Functional brain imaging studies of cortical mechanisms for memory. *Science, 270*, 769-775.

Vallée-Tourangeau, F., Anthony, S. H., & Austin, N. G. (1998). Strategies for generating multiple instances of common and ad hoc categories. *Memory, 6*, 555-592.

Vanderplas, J. M., & Garvin, E. A. (1959). The association value of random shapes. *Journal of Experimental Psychology, 57*, 147-154.

Von Stein, A., Rappelsberger, P., Sarnthein, J., & Petsche, H. (1999). Synchronization between temporal and parietal cortex during multimodal object processing in man. *Cerebral Cortex, 9,* 137-150.

Vosniadou, S., & Ortony, A. (1983). The emergence of the literal-metaphorical-analogous distinction in young children. *Child Development, 54,* 154-161.

Walker, W. H. & Kintsch, W. (1985). Automatic and strategic aspects of knowledge retrieval. *Cognitive Science, 9,* 261-283.

Weisberg, R. W. (1986). *Creativity: Genius and other myths.* New York: W. H. Freeman and Company.

Williams, M. D., & Hollan, J. (1981). The process of retrieval from very long-term memory. *Cognitive Science, 5,* 87-119.

Williams, M. D., & Santos-Williams, S. (1980). Method for exploring retrieval processes using verbal protocols. In R. Nickerson, (Ed.), *Attention and Performance, VIII,* (pp. 671-689). Hillsdale, NJ: Lawrence Erlbaum.

Wingfield, A., Lindfield, K. C., & Kahana, M. J. (1998). Adult age differences in the temporal characteristics of category free recall. *Psychology and Aging, 13,* 256-266.

Winner, E., McCarthy, M., & Gardner, H. (1980). The ontogenesis of metaphor. In R. P. Honeck and R. R. Hoffman, (Eds.), *Cognition and figurative language* (pp. 341-361). Hillsdale, NJ: Lawrence Erlbaum.

Wittgenstein, L. (1973). *Philosophical investigations* (3rd ed.) (G. E. M. Anscombe, Trans.). Engelwood Cliffs, NJ: Prentice Hall. (Original work published 1953).

Wixted, J. T., & Rohrer, D. (1994). Analyzing the dynamics of free recall: An integrative review of the empirical literature. *Psychonomic Bulletin & Review, 1,* 89-106.

Wood, G., & Underwood, B. J. (1967). Implicit responses and conceptual similarity. *Journal of Verbal Learning and Verbal Behavior, 6,* 1-10.

Wu, L., & Barsalou, L. W. (2003). Perceptual simulation in property generation. Manuscript under review.

Yumino, K. (1977). An analysis of the retrieval process in long-term memory by free emission method. *Japanese Journal of Psychology, 48,* 7-13. [Abstract]

Zelterman, D. (1999). *Models for discrete data.* New York: Oxford Press.

Zimmer, H. D. (1989). Visuelle und semantische Merkmale im Recall [Visual and semantic characteristics in recall]. *Sprache und Kognition, 8,* 115-125.

Zimmer, H. D., & Engelkamp, J. (1989). Does motor encoding enhance relational information? *Psychological Research, 51,* 158-167.

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science, 13,* 168-171.

Notes for Chapters 1 and 2

[1] In many of the sciences, the term morphology is often used to refer to the shape, form, or structure of some thing or pattern, or to the study thereof, within a science. Also note that by the use of the words "shape" and "form" in this thesis, I am generally referring to three-dimensional object shape, unless specified otherwise.

[2] It is not difficult to imagine that similarity organization in isolation, if unconstrained, could easily become maladaptive under many circumstances. I suggest that the tendency is often constrained by the fact that objects in the real world consist of correlated features, such that we are likely to be reminded of one thing by another that overlaps with it in a combination of multiple respects. Of course, the similarity-based tendencies are also guided by such factors as context and the person's goals, which help ensure relevance.

[3] Young children can also use functional analogies that are not rigidly constrained by shape. For example, a toddler may sit on a beachball, calling it a "chair" (Anglin, 1983), even though the spherical shape of the object does not optimally afford sitting. This indicates that young children can be flexible in their analogical thinking. Nevertheless, it remains true that shape-based analogies are common in toddlers' language and behaviour.

[4] Colour, surface patterns, and textures do provide important additional information about an object, particularly in combination with shape information. For example, colour often signals fruit or vegetable ripeness. However, the elements of many patterns—spots and stripes, and so on—are themselves shapes or elements of shapes such as contours and angles. Much the same could be said of visible textures such as mammalian fur, fish scales, bird feathers, and such. Removing those shapes or shape elements would eliminate the specific entity information provided by the pattern.

[5] By supramodal, I refer to information that (i) can be registered through more than one sensory/perceptual modality, such as time, space, intensity, and so on, or (ii) emerges out of the synthesis of unimodal sources (e.g., flavour perception). Some authors (e.g., Smith, 1987) refer to such information, particularly (i), as amodal. However, supramodal information is specific or limited to certain modalities. For example, spatial information arises from visual, auditory, and various somatosenses, and is probably also often encoded in conjunction with action information, but does not seem to arise from smell and taste. Moreover, smell can combine with taste information in a way that other sensory information cannot normally combine with either taste or smell.

[6] Usage of the term *perceptual* can be confusing because, in perceptual theories, motor, linguistic, bodily state, emotional, social, abstract conceptual, introspective, and cognitive operational information are all included (Barsalou, 1999). Another point of clarification here is that perceptual representations are not strictly analogical or non-arbitrary. For example, words are perceptuomotor representations that are normally related arbitrarily to their referents. Rather than introduce a new term, I will still use perceptual here, despite the above caveats and clarifications, mainly because (a) the term has emerged in

predominant usage in recent theoretical discussions, (b) the type of information studied here involves perceptual properties (i.e., shape), and (c) the shape representations activated during FE and FR are hypothesized to be implemented in perceptual areas of cortex.

[7] The Bousfield et al. (1959) study involving geometric figural similarities is of some concern. Other researchers have previously described its clustering results as negligible, requiring "several trials" to occur (Engelkamp & Zimmer, 1994; Frost, 1972). That description needs to be corrected. In fact, moderate-to-strong clustering occurred on every trial except Trial 1. There were 24 abstract geometric figures to recall, with six exemplars in each of four geometric design categories. The mean Ratio of Repetition (RR) clustering scores (where chance = .2173) for Trials 1, 2, 3, 4, and 5, respectively were .22, .41, .49, .55, and .60. The fact that the mean clustering for Trial 1 approximates chance level (.217) must be regarded with due caution, because (a) there were no practice trials; (b) learning the task itself was "difficult" (Bousfield et al., 1959, p. 284)—probably most difficult on Trial 1; (c) the stimuli were unfamiliar, non-meaningful, and in many cases highly confusable; and (d) there was very low mean correct recall (5.76 / 24 figures) on Trial 1. The second trial mean RR of .41 is consistent with significantly-above-chance clustering found for moderate semantic category clustering (e.g., Bousfield, Cohen, & Whitmarsh, 1958). Taking all of these facts into consideration, there is good support for the original conclusion of Bousfield et al. (1959) that geometrically-based categorical organization can occur in the learning of new design patterns.

[8] The authors of both of these studies reported variation or elaboration on themes in sequential productions of non-meaningful geometric designs (Jones-Gotman & Milner, 1977) and gestures (Jason, 1985) by healthy participants. In other words, there were clusters, in which elements of one production were carried over into the next.

[9] The numbers of participants who received flat first versus long first were not as evenly balanced as they could have been (i.e., long was first in 4 cases and flat was first in 7 cases).

[10] In the pilot phase, before an exact time limit had been decided upon, 2 pilot participants did the FE task for approximately 6 1/2 minutes. Their production rates slowed considerably toward the end of that time period, so the number of responses produced would not have differed much if they had completed a full 7-minute task.

[11] Arguing from simulation results, Brook and Stirling (1984) suggested that the chi-squared distribution was more sensitive than Brennan and Light's (1974) normal Z distribution under certain circumstances. The present investigation, with real data, suggested that the Brennan and Light approach was slightly more sensitive overall. This may have been owing to the fact that, compared to Brook's and Stirling's (1984) r x c tables, those in the present research were very large and sparse, characteristics that could make the Z-score approximation more appropriate than the chi-squared (Morris, 1975, cited in Zelterman, 1999). Nonetheless, because the chi-squared was more convenient to compute, it was used for most of the analyses of agreement here.

[12] In calculating the η statistic, size was treated as an interval variable, while shape and semantic variables were treated as nominal. This was calculated using the basic rows x columns Crosstabs procedure in SPSS 13.0, GradPack. Although size in the present study would more accurately be described as an ordinal variable, nominal-by-ordinal statistics were not available for these calculations. Yet, because size is an ordered variable here, treating size as an interval variable probably provides a more accurate measure of agreement than would be provided by the chi-squared. In any case, ordinary chi-squared results also showed that the size-semantic category and size-shape category agreement levels were significant, comparable to levels reported for shape-semantic category.

[13] Indeed, cross-correlational analyses, not reported here, suggested that shape similarity was related to IRT durations not only at lag 0 (where the relation was strongest), but also (with diminishing levels of association with increasing lag or lead) at lags and leads of 1 for most domains, and up to two or three for the round domain.

[14] In her 1971 article, Frost often uses the word "shape," though her experimental categories are by orientation.

[15] Frost (1971) also employed a control group (n = 15) that received only the object names in both practice and test trials. For that group, clustering that was scored according to experimenter-defined orientation categories was below chance level (M < .25). Therefore, it is unlikely that Frost's depicted orientation information might have in some way been confounded with the object names. These control participants later reported, in the post task interview, what orientation categories (out of the four experimenter-defined categories) they had imaged or else would have expected for each item. Frost (1971) then rescored their recall outputs according to those subjective orientation categories. Interestingly, above-chance clustering was obtained (M = .419, where chance = .25) according to those subjective orientation schemes, which suggested to Frost (1971) that there may have been an imaged orientation effect even for the participants who were exposed only to the names of the objects for study and recall.

Engelkamp and Zimmer (1994) claimed that Frost's (1971) significant orientation clustering depended on a practice trial in which participants had to recognize object pictures (a "practice picture" (PP) procedure). However, there is evidence in Frost's (1971) study showing that the PP procedure was not necessary to produce the clustering effect. Another group of participants received a "practice word" recognition (PW) condition. In the PW condition, participants were shown pictures of the objects in the practice study, but then in the practice test they were shown only the names of objects and were required to recognize the names of those objects that they'd seen in the practice test. Thus, these PW participants were prepared for a test of recognizing objects by name, not by visible cues such as depicted orientation. This PW group later received the regular study-recall test trial where the object picture stimuli were presented. The PW group had clustering scores on their test trial (M_{PW} = .506) that were as high as for two other groups (M_{PP1} = .525 and M_{PP2} = .439) that had earlier received the "practice picture" (PP) trial. In a direct contrast between M_{PW} and M_{PP1}, $F < 1$. Therefore, the practice picture procedure was not a critical factor for the clustering effect.

Engelkamp and Zimmer (1994) also suggested that the significant orientation clustering in Frost's studies may have been due to the allegedly high salience of the orientation information in the to-be-recalled stimuli. Although this is a reasonable suggestion, Frost (1971) reports that, in post-task interviews where the participants were asked about their "...methods of retention and retrieval" (p. 410), *none* of the 45 participants exposed to the picture stimuli reported any awareness of the orientation categories (p. 413). It is possible that they could have been aware of the categories, but, for unknown reasons, did not report it. However, if orientation information was as salient as Engelkamp and Zimmer suggested, it should have been reported by some participants.

[16] Some of Zimmer's mean ARC scores (.13 and .12) may have been significantly greater than chance (zero), but the measures of variability needed to assess this possibility were no longer available (H. D. Zimmer, personal communication, November, 2001).

[17] The 2 participants whose data were excluded had the lowest and second lowest mean recall in the study, though there were additional reasons for exclusion. One participant misunderstood the meaning of two of the words, such that both shape and semantic category were changed drastically for both words (no other participants misunderstood those two words); reported a large number of repeated items; included irrelevant speech during recall; had very low recall on the first two trials (four and seven items recalled correctly on Trials 1 and 2, respectively; mean recall across all trials = 8.75); and grouped only four items (two units of Size 2) in the sort. Note that the significant misconstruals of two words *and* the large number of repeats were deemed sufficient to warrant exclusion. The other participant reported that he had experienced a stressful negative event prior to the session, and that he believed that this had affected his recall performance; and he had very low recall on the first three trials (mean 7.3 items recalled over the first three trials, with a mean of 8.5 over all trials), included irrelevant speech during recall, had considerable difficulty with the backwards counting task, and grouped only two items in the sort. For this latter participant, emotional state, coupled with low recall, was deemed sufficient to warrant exclusion.

[18] Two other participants had been exposed to fragments of information that might have been of concern. One had heard about a "round things" task, and another had done a mental imagery task involving the combination of different shapes. Neither of these participants appeared to have been influenced in a way that would increase shape clustering—that is, their mean clustering scores were around chance, and both had first trial clustering scores below chance. Therefore, the data from these participants were included in all analyses.

[19] In the tests of ARC and MRR clustering, the chance level of clustering constitutes the null hypothesis criterion to be rejected or retained. Therefore, directional (one-tailed) tests were used, because the upper tail of the distribution of clustering scores was irrelevant in this case. I am also in agreement with Hays (1994, see pp. 297-299) that a directional test is appropriate when the researcher has an a priori directional hypothesis and empirical and theoretical reasons motivating that hypothesis. Unless noted otherwise, directional tests were used in all of the directional, a priori hypothesis tests, and non-directional tests were

used in post hoc tests in this study.

[20] There were 13 alternative items that were used for the 2 of the 33 participants who received the first version of List 3. Because there were inadequate numbers of IRTs from which to derive valid means for those items, approximations were obtained by calculating the mean cluster IRT and mean switch IRT from the pool of 13 alternative items, for both beginning- and end-word analyses. This amalgamated case was then added to the main $n = 60$ items analyses, resulting in $n = 61$ cases. For the amalgamated beginning word, cluster $M = 2.128$ s ($n = 16$) and the switch $M = 2.320$ s ($n = 58$). For the amalgamated end-word, cluster $M = 2.665$ s ($n = 14$) and the switch $M = 1.989$ s ($n = 57$). The results for the beginning- and end-word analyses that included the amalgamated item are as follows: for the beginning-word, $t(60) = 4.74$, $SD = 0.967$, $p = .000007$ (one-tailed), $\eta^2 = .27$; and for the end-word, $t(60) = 4.31$, $SD = 1.124$, $p = .00003$ (one-tailed), $\eta^2 = .24$. These results are nearly the same as those for the main $n = 60$ analyses.

[21] There are several other methodological differences between Zimmer's (1989) FR study and the present one that could help explain the differing results. However, a thorough methodological comparison is not feasible at this point because Zimmer's (1989, Exps. 1-3) object lists were not published and are not available (H. D. Zimmer, personal correspondence, November, 2001).

Chapter 3. Toward a Neural Theory of Shape Categorical Clustering in Free Emission and Free Recall Retrieval

According to perceptual theories (Barsalou, 1999; Fuster, 1995; Hebb, 1949, 1968; Martin, Ungerleider, & Haxby, 2000), when we retrieve an object noun from LTM during free emission (FE) or free recall (FR), some perceptual properties of the referent will tend to be reactivated, including its form. The neural assembly for a snake's form, for example, shares some form-selective neurons with those of other similarly-shaped entities (e.g., rope, hose, etc.). This overlap of shape-selective neurons between similar object representations is posited to be the primary facilitator of object shape clustering and the accompanying cluster-IRT reductions. In this chapter, I will attempt to describe in more detail the hypothesized mechanisms involved in this overlap/similarity-based clustering.

Hebb (1949) was one of the first to present explicitly the idea that neural assembly overlap could play a role in the organization of recall. He speculated that if the neural assemblies of items A and B overlapped on a common assembly C, then activation of A could promptly facilitate the activation of B, through the persistent activation of C. He wrote: "...the subject associates an object B with an object A, because both are associated with something else, C: a familiar trick in memorizing lists of words or nonsense syllables in learning experiments." (1949, p. 131). Hebb (1949) also implied that A and B could be object assemblies, and C could be a shape assembly common to both objects: "A perceived object consists of a number of perceptual elements...The same elements recur in different perceptions, so that two concepts to be associated may have phases (assembly actions) in common" (p. 130). The object assemblies would be connected to, and would activate, their respective word assemblies.

One of the few neurally-based explanations of clustering in recall tasks was put forth by Bousfield and Cohen (1953). Their brief explanation, which focused on semantic categorical clustering, also drew upon Hebb's (1949) notion of the "overlap of assembly action" (Bousfield, 1953, p. 239) between items that share a superordinate (e.g., animals).

However, as was discussed in chapter 1, demonstrating purely perceptual clustering of concrete entity/object nouns presents certain problems. In the on-line task performance context of FE or FR, object activations are probably too brief and variable to reliably engage most of the constituent perceptual unit assemblies of the object assemblies. Thus, it cannot be guaranteed that most perceptual unit assemblies will be significantly activated each time their object assembly is activated. In a statement relevant to this point, Hebb noted that "...a concept is not unitary. Its content varies from one time to another, except for a central core whose activity may dominate in arousing the system (i.e., the concept's neural assembly) as a whole." (1949, p. 133, parentheses added). When a concept is activated, its "dominant core" should be at least partially activated, including under conditions of "limited stimulation" (p. 133). I suggest that, in light of the importance of object shape in object representation (see chapter 1), a dominant core of an object assembly in LTM would generally be its shape assembly. Hebb (1949) believed that the links between object assemblies via a common perceptual unit assembly (e.g., shape) are established prior to the experimental situation. If the object assemblies are already established with their shape assembly in LTM, they need only be activated sufficiently in the experimental context to produce clustering effects.

Although the theory described herein was partly inspired by Hebb's (1949) neural assembly theory, several fundamental changes were made in light of what is now known about chemical transmission, inhibition, non-Hebbian learning, and so on (for further discussion, see Milner, 1999). My goal is to develop a principled neurally-based theory that is applicable to a wide variety of cognitive tasks. A sub-goal is to explain clustering in retrieval sequences. This chapter focuses on shape categorical clustering as a primary example, but will touch upon alternative clustering schemes, such as spatial relational and semantic categorical, that commonly arise. I will present here the abbreviated versions of some of the assumptions, mechanisms, and principles by which clustering is regulated.

3.1. General Perceptual-Theoretic Principles

3.1.1. Reactivation

Perhaps the most important underlying assumption of perceptual theories of memory is that a subset of the perceptual and motor neural populations that are activated during perception and interaction with an object are reactivated during retrieval of that information about that object (see below). Critically, this reactivation can occur in the absence of the original stimulus object. A wide range of evidence supports this assumption. Reactivation of perceptual (and motor) association cortices during retrieval of concrete entities, or of perceptual information pertaining to those entities, has been shown in ERP and EEG studies (Heil, Rösler, & Hennighausen, 1996; Rösler, Heil, & Hennighausen, 1995), in neuroimaging studies (for reviews, see Martin, 2001, 2007; and see Handy et al., 2004; Moscovitch, Kapur, Köhler, & Houle, 1995; Nyberg, Habib, McIntosh, & Tulving, 2000; Roland & Friberg, 1985; Thompson-Schill, Aguirre, D'Esposito, & Farah 1999; Vaidya, Zhao, Desmond, & Gabrieli, 2002; Vanderberghe, Price, Wise, Josephs, & Frackowiak, 1996; Wheeler, Petersen, & Buckner, 2000), and in studies using direct microelectrode stimulation of human cortex (Penfield & Perot, 1963; Penfield & Rasmussen, 1950). Lesions to perceptual association cortices generally result in partial to complete loss of the ability to retrieve certain perceptual information, with the type of deficit depending on the location of the lesion (e.g., Levine, Warach, & Farah, 1985; Warrington & McCarthy, 1987; for reviews, see Damasio, 1989; Farah, 1990; Humphreys & Forde, 2001). Reactivation of appropriate perceptual (and/or motoric) cortices during imagery experience occurs in at least the visual (Kosslyn, 1994), motor and sensorimotor (Ehrsson, Geyer, & Naito, 2003; Hauk, Johnsrude, & Pulvermüller, 2004; Jeannerod, 1994), tactile (Yoo, Freeman, McCarthy, & Jolesz, 2003), auditory (Zatorre, Halpern, Perry, Meyer, & Evans, 1996; Zatorre & Halpern, 2005), and gustatory (Simmons, Martin, & Barsalou, 2005) modalities. Direct microelectrode recordings from human medial temporal areas show that neurons firing selectively for a specific stimulus (e.g., a picture of a face, animal, food item, household object, etc.) during perception are reactivated during mental imagery of that stimulus (Kreiman, G., Koch, C., & Fried, I., 2000a, 2000b). Thus, a wide variety of evidence supports the principle of reactivation.

Particularly compelling evidence of the importance of reactivation in retrieval and clustering comes from a recall study that used fMRI (Polyn, Natu, Cohen, & Norman, 2005). Participants had to recall the photos of famous locations, famous faces, and objects presented during the encoding phase. Participants tended to recall items when there was a high degree of reinstatement of the appropriate category-specific pattern of brain activity. For example, if the pattern of activity that occurred while encoding faces reoccurs to a high degree at some point in retrieval, recall of faces occurs at that point.

The recall sequences included clustering, wherein the reactivation of category leads to recall of items from that category; when another category is reactivated, then items are recalled from that one, and so on. As Polyn et al. (2005) wrote, "As subjects search memories from a particular event, their brain state progressively comes to resemble their brain state during the sought-after event, and the degree of match predicts what kinds of information they will retrieve." (p. 1966).

Retrieval in FE or FR involves more than simply the reactivation of some neural ensembles that were active in previous experiences. As Peter Milner (1999) has pointed out in discussing stimulus recognition, some kind of "familiarity engram" (p. 85) must be activated in conjunction with the reactivation of the target stimulus' neural ensemble(s) in order for recognition to be successful. Similarly, in recall, one must reactivate not only the previously-studied stimuli, but also, in conjunction with those stimuli, the task episode in which they were presented.

As Barsalou (1999) has proposed, the contents of reactivation are partial and sketchy, not complete reproductions of all details of a target entity. This is likely the case in FR and FE, where only a very brief and partial reactivation of the concrete information pertaining to an entity (e.g., concept of an apple) is needed to activate the appropriate lexical representation (e.g., the word apple).

There can be reactivation of perceptual areas without the activation of any memory per se. For instance, one could imaginally arrange a new abstract composition of lines and colours, without the image representing anything familiar. More generally, reactivation of perceptual and motor areas is likely to occur in a variety of ordinary instances of productive thinking, such as in planning any behaviour. In these cases, elements are reactivated but combined in new ways.

3.1.2. Neuronal Selectivities

A critical assumption in the present theory is that neurons involved in perceiving and remembering entities/objects have stimulus selectivity. Individual neurons in perceptual cortex are selective for (or are "tuned" for) specific features and feature combinations, as has been demonstrated in primate inferior temporal (IT) cortex in regard to shape features (for a review, see Tanaka, 1996). For example, neuron A may dramatically increase its firing rate to oval stimuli, whereas its firing rate may remain at baseline level or lower for square stimuli. Likewise, neuron B may respond to a red, but not green, colour. Neuron C may respond most strongly to red oval stimuli. In addition, visual and visual-tactile neurons that are selective to the three-dimensional shapes of objects (e.g., flat rectangles, long cylinders, rings, etc.) occur in primate parietal cortex (Sakata et al., 1999), as do neurons selective for two-dimensional shapes (Sereno & Maunsell, 1998). An analogous region selective for shape has been identified in humans (Oliver & Thompson-Schill, 2003).

Neurons that have selectivities that are similar to each other tend to be clustered together in microcolumns perpendicular to the cortical surface (Tanaka, 1996; Wang, Tanaka, & Tanifuji, 1996). Neurons selective for visual stimuli (e.g., faces and objects) have also been located in monkey prefrontal cortex (PFC) (e.g., Ó Scalaidhe, Wilson, & Goldman-Rakic, 1999). Some stimulus-selective neurons remain active when a stimulus representation is maintained over a delay period wherein the stimulus is no longer present; such neurons are found in IT (Fuster & Jervey, 1981; Sakai & Miyashita, 1991), posterior parietal (PP) (Andersen, Bracewell, Barash, Gnadt, & Fogassi, 1990), and in

PFC (Desimone, 1996). Different PFC neurons may show stimulus cue-selective delay activity for faces or objects (Ó Scalaidhe et al., 1999), colour (Quintana, Yejeya, & Fuster, 1988), tactile feature (Romo, Brody, Hernández, & Lemus, 1999), acoustic tone (Bodner, Kroger, & Fuster, 1996), spatial position (Courtney, Petit, Maisog, Ungerleider, & Haxby, 1998; Funahashi, Bruce, & Goldman-Rakic, 1989; also see the fMRI study by Leung, Gore, & Goldman-Rakic, 2002), and the conjunction of spatial position and object (Rao, Rainer, & Miller, 1997). Some neurons in lateral PFC (lPFC) appear to be tuned in a manner that suggests sharp discrete categorical boundaries between complex three-dimensional animal shapes (e.g., some neurons are, or through training can be, selective for cats, or else for dogs; see Freedman, Riesenhuber, Poggio, & Miller, 2001).

Regarding action-sequence selectivities, Tanji (2001) and others have shown that many neurons in the pre-supplementary motor area (pre-SMA) in monkey cortex are selective for action sequences per se. That is, if neuron *j* responds to a specific cue for preparing action sequence *a-b-c*, it will not respond to cues for any other ordering of those same action components, such as *a-c-b*, etc.).

Individual objects are not represented in a one-to-one manner by individual neocortical neurons (i.e., "grandmother cells"), nor are objects represented in a purely distributed manner involving all neocortical neurons. Rather, the neuroscientific evidence strongly suggests that objects are represented by constituent populations of neurons in multiple regions, at multiple levels, activated in combination (for discussions and reviews, see Damasio, 1989; Fuster, 1995; Tanaka, 1996; and for empirical evidence, see Haxby et al., 2001; Ishai, Ungerleider, Martin, Schouten, & Haxby, 1999; Ishai, Ungerleider, & Haxby, 2000). I propose there is population coding at the level of individual features (e.g., shape, colour, hardness, taste) of an object, and at the level of coherent object representation, where the constituent feature populations combine to form larger, more diverse populations.

3.2. Neural Assemblies

Hebb (1949, 1968) and many subsequent researchers (e.g., Braitenberg, 1978; Damasio, 1989; Fuster, 1995; Pulvermüller, 1999) conceived of a neural assembly primarily as a group of cells that were mutually excitatory. I believe evidence and logical considerations now show that inhibitory neurons also play important roles in neural assemblies (also see Stent, 1973).

First, although pyramidal cells in neocortex generally outnumber inhibitory interneurons by about 4 to 1, inhibitory neurons do exist in cortex, and they are activated by excitatory inputs. Those inhibitory neurons then send inhibitory input to excitatory and inhibitory cells. When pyramidal cells increase their firing rate, so do many of the inhibitory neurons that receive inputs from them. Indeed, some inhibitory neurons that use electrotonic synapses, such as basket cells, would activate almost instantly after their neighbouring connected pyramidals were activated.

Second, I conceive of neural assemblies as extending beyond cortex to incorporate subcortical areas, and this means that inhibitory neurons must be taken into account. For example, most of the basal ganglia (BG) consist of inhibitory neurons (except for the subthalamic nucleus, which contains excitatory neurons), and the BG is critical in action sequencing and possibly other types of representation (Levy, Friedman, Davachi, & Goldman-Rakic, 1997; Miller & Wickens, 2000).

Third, it is difficult to conceive of how conjunctive (or conditional, combinatorial) neurons could in fact be conjunctive without some inhibition of at least some excluded features. For example, a neuron highly selective for stimulus AB may not be very responsive to A, B, AC, BD, and so on. A purely excitatory scheme cannot readily explain this type of selectivity, whereas the inclusion of inhibitory connections linking the AB-selective neurons to A, B, C and D, and conjunctive assemblies such as AC and BD, provides the beginning of an explanation for the selectivity.

Fourth, blocking the effects of inhibitory neurons disrupts or alters the tuning characteristics of other shape-selective neurons nearby (Wang, Fujita, & Muryama, 2000). Thus, inhibitory neurons play an important role in regulating the tuning characteristics of the neurons to which they are connected, and therefore must be taken into account in an explanation of the stimulus selectivity of neurons and neural assemblies.

Most theorists, including Hebb (1949), postulated two main types of assemblies, which can be called active and stable. The stable assembly consists of preestablished enduring synaptic connections between neurons, such that certain interconnected neurons become activated together in combination to the exclusion of other neurons. I propose that the stable type of neural assembly is an established organized binding of excitatory and inhibitory connections between neurons that are activated to represent (a) a simple feature, (b) a feature conjunction (e.g., a simple shape), (c) a simple object, (d) a recurring configuration or constellation of components (e.g., layout of a familiar room, a complex object such as a car), (e) simple action, and (f) syntactic ordering of actions (for discussions regarding the possible schemes for the implementation of syntactic ordering, see Lashley, 1951; Rizzolatti, Fogassi, & Gallese, 2000; Rizzolatti & Arbib 1998; Tanji, 2001). Complex widely distributed assemblies such as event types are composed of the elemental (a)-(f) assemblies. The "bindings" and "connections" referred to above are working synaptic connections that are established through development and experience.

The process by which these stable assemblies are established (or consolidated) is generally assumed to involve interaction between medial temporal lobe (hippocampus and immediately-surrounding cortex) and neocortex. At the cellular level, long-term potentiation (LTP) and long-term depression (LTD) are believed to be important in helping to establish stable assemblies.

Episodic memory is not discussed in detail in this chapter, but I wish to add that first-person experience and its reactivation involves binding of elemental assemblies including (a)-(f) with elemental bodily state assemblies and diffuse higher-order bodily-state assemblies that contribute to first-person self-reference. Thus, if a person has experienced stimulus X, to remember X as an experience requires that X be bound to bodily state assemblies, as well as spatio-temporal contextual assemblies, provided that there was sufficient encoding to establish the binding, and that the assemblies are activated in combination at retrieval. Established assemblies are the bases for long-term memories.

Established (i.e., stable) assemblies must be activated, of course, during encoding and retrieval. Anatomically, active and stable assemblies occur in–or at least overlap on–the same neural substrate (Fuster, 1995). An active assembly is dynamic and flexible, and is not confined only to memory processes per se, though the format is also constrained by the more stable, LTM assemblies. These types of working memory (WM) and STM assemblies implement the general spatio-temporal representational frameworks for normal thinking, reasoning, planning, problem solving, retrieval operations, and creative thinking. These flexible assemblies underlie the various cognitive operations used in FR and FE. The initial establishment of the LTM assemblies also depends on the configuration of the active type during the encoding stage. The active assemblies that implement WM and STM themselves can become established with experience.

3.3. General Assumptions

1. Excitatory and inhibitory shape-selective neurons are distributed predominantly in secondary occipital, inferior temporal (IT), parietal (PC), and prefrontal (PFC) cortices. Neurons with similar shape selectivities, in these different regions are connected to each other by long fast-conduction myelinated axons. Within these respective areas, there are also white-fibre connections, and intracortical (mostly non-myelinated) connections between shape-selective neurons.

2. Shape assemblies consist of shape-selective neurons bound through synaptic connections.

3. There are assemblies for words, objects/entities, actions, spatial contexts, task sets, episodes, motivations, bodily states, and other constructs.

4. The synaptic connections within and between assemblies are established, and can change, due to experiential and genetic influences.

5. Assemblies often occur within larger assemblies and contain smaller constituent assemblies (subassemblies).

6. Assemblies can share membership in larger assemblies and can overlap (share

constituent neurons) with other assemblies on their respective subassemblies.

7. The greater the similarity between two shape representations, the greater the overlap of the neural constituents between them.

8. Assemblies that preserve LTM representations are flexible and dynamic, yet have dominant cores of relatively enduring synaptic connections.

9. At any given time, an arrangement of neural assemblies is active in working memory.

3.4. Basic Mechanisms and Constraints

A currently-active object assembly can facilitate (ease, speed up) the activation of another object assembly that has a similar shape, due in part to overlap of the respective shape assemblies of the object assemblies. What mechanisms could plausibly support the immediate successive retrieval of same-shape items?

3.4.1. Sustained Neuronal Activity

Facilitated activation of one object assembly by the activation of another can occur through sustained, relatively elevated firing (in the time frame of a few seconds to 15-30 s) of the shape assembly neurons that overlap in the respective object assemblies' neurons in lateral PFC, posterior neocortex (PNC; here refers chiefly to temporal and parietal associative neocortex), and medial temporal lobe (MTL; chiefly entorhinal cortical [e.g., see Egorov, Hamam, Fransen, Hasselmo, & Alonso (2002)] and possibly hippocampal neurons). The nature of the sustained activity differs, in some respects, from region to region (those regions being PFC, PNC, and MTL). The key difference between the sustained activity of PFC and PNC neurons is that, in tasks such as delayed matching-to-sample (DMS), the PFC neurons in question will sustain their mnemonic activity robustly for a stimulus cue A during a delay in which other different stimuli B, C, D, and so on, are presented subsequently, whereas it is usually the case that the sustained mnemonic activity of the PNC neurons for A is disrupted and curtailed when B, C, D, and so forth, are presented subsequently (Desimone, 1996). I posit that the distracter-resistant PFC sustained activation is one of the main mechanisms through which same-category items are clustered during FR list encoding, as well as during retrieval in either FR or FE.

Interestingly, the mechanism of sustained activity in clustering effects has been conceived in recent years by multiple authors, apparently independently of one another. In 2001, I suggested a primary role for sustained activation in FR and FE clustering. Subsequently, I found that Rolls and Treves (1998) in their modelling of the fluency task (i.e., FE) and Becker and Lim (2003) in their modelling of FR, had also included a role for sustained activity in clustering effects. Of course, Hebb (1949) appears to have been the first to have proposed it.

3.4.2. Short-Term Synaptic Plasticity

Short-term synaptic enhancement (STSE; for a review, see Fisher, Fischer, & Carew, 1997) and depression (STSD; see Schneggenburger, Sakaba, & Neher, 2002), operate in the time frame of 10s of milliseconds to 10s of seconds. Whereas STSE is believed to involve an increased influx of calcium ions into the presynaptic terminal, resulting in increased amounts of neurotransmitter being released and enhanced post-synaptic potentials, STSD is believed to involve a reduced influx of calcium ions into the presynaptic terminal, resulting in reduced neurotransmitter release and reduced post-synaptic potentials. There will tend to be STSE for an object's shape assembly activated at time 1, accompanied by STSD for concurrently less-activated other-shape object assemblies. This will bias activation at time 2 (and 3, and 4, etc.) in favour of those shape assemblies having the most overlap with the shape assembly activated at time 1.

Although STSE can result from or lead to sustained activity, I propose that some clustering can also occur simply on the basis of STSE mechanisms with no involvement of sustained activity. For example, priming may occur in retrieval such that the prior retrieval of an object item of shape A increases the probability that some other object item having shape A will be activated sooner than object items having shape B, or C, or D, and so on.

3.4.3. Longer-Term Synaptic Plasticity

Short-term potentiation (STP) and depression (STD), lasting a few minutes to tens of minutes, and long-term potentiation (LTP) and depression (LTD), lasting hours to days, can result in an increase in the potential for activation of the shape assemblies involved in clustering, in addition to shorter term enhancement/depression that would occur during the early stages of their induction. Note that the depression (STD or LTD) aspect would tend to reduce the probability of clustering, for example, on later trials of FR, for available assemblies that did not get used for clustering in the initial trials of FR. Longer-term synaptic plasticity should contribute to increases in clustering with successive recall trials. Note that LTP involves intracellular protein synthesis, contributing to preestablished enduring synaptic connections (see below), whereas LTD involves intracellular protein breakdown, leading to the weakening of synaptic connections.

3.4.4. Other Dynamic Factors

Numerous additional short-term dynamic factors affect synaptic transmission, including changes in extracellular currents; depletion of neurotransmitter; hemodynamic factors such as changes in blood flow velocity to specific regions and dilation or constriction in the microscopic blood vessels supplying the active neurons; and so on. Given that an object assembly has just been activated, these factors could help enhance the next similar object assembly's activation, within roughly the same time frame as short-term synaptic plasticity (10s of milliseconds to 10s of seconds).

3.4.5. Preestablished, Enduring Synaptic Connection Strengths

Sustained activity and short-term synaptic plasticity must be constrained by preestablished, enduring synaptic connections between an object assembly and its shape assembly, and between shape assemblies. Otherwise, sustained activity and short-term synaptic plasticity might facilitate clustering for whatever overlapping subassembly activations happened to coincide between object assemblies at any given moment in time, likely leading to haphazard, unstable clustering with respect to any given subassembly type. Because the basic schematic shape of an object is perceived almost every time an object is encountered, strong synaptic connections should become established for the shape assembly within the object assembly. The shape assembly should form a dominant core assembly of the object assembly. In addition, there should be strong connections within the shape assembly, and between highly overlapping (i.e., highly similar) shape assemblies.

Preestablished enduring synaptic connection strengths are the result of long-term chemical, electrical, and structural changes in the pre- and post-synaptic neurons within and between assemblies. The structural establishments include growth in the size and number of the axons, dendrites, and terminals of the neurons so affected (Huttenlocher, 2002; Rosenzweig, Leiman, & Breedlove, 1999). (Note that there is also a subtractive process during development involving pruning of unused synapses). The functional establishments include a reliable, lasting increase (or decrease) in synaptic transmission, more (or less) neurotransmitter release, larger (or smaller) post-synaptic potentials, etc. Once these long-term structural and functional modifications have been established, there will be a tendency for similar-shape (i.e., overlapping) object assemblies A and B to be activated closely in succession, even though the full object assemblies A and B may have never been activated contiguously before (as implied by Hebb, 1949, see p. 132).

Numerous structural factors affect synaptic transmission, including differences in neuron type, synapse type, myelination versus non-myelination of axons and dendrites, effects of glial cells in synaptic actions, relative numbers of established excitatory and inhibitory inputs onto a neuron, interneuronal distances [e.g., neurons with similar shape-selectivities tend to be clustered together in thousands of microscopic columnar clumps (Tanaka, 1996, 2002); reduced interneuronal distances reduce the transmission times between them], and so on. All assemblies that implement long-term memories depend on preestablished connection strengths, as does much of the structure of working memory.

3.4.6. Motivation-Related Effects on Synaptic Transmission

The activation of motivational assemblies, in conjunction with task-set assemblies (see below), is important in initiation, termination, maintenance, and switching of various assembly activations. For discussions of and evidence for the neural systems supporting motivation, see Hebb (1949); Fellows and Farah (2003); O'Doherty, Critchley, Deichman, and Dolan (2003); Sakagami and Watanabe (2007); Samejima and Doya (2007); and Wallis and Miller (2003). (Note that those various authors do not literally use the phrase "motivational assemblies"). Some motivational assemblies signal "reward" (or "do-it-again" appetitive states) and others signal "punishment" (or "stop-doing-it" aversive states) to the task-set assemblies and their subassemblies that are engaged during performance. These motivational assemblies modulate sustained activity and synaptic plasticity processes by either enhancing or depressing synaptic transmission for the various presently-active assemblies, usually by the release of neurotransmitters and neuromodulators. Motivational assemblies will tend to enhance activation for quick and easy task-set assembly and object assembly activations, and will tend to depress temporarily the activation for slow and difficult task-set and object assemblies. Motivational assemblies span numerous cortical and subcortical structures, but, for FR and FE performance, these especially include the ventral tegmental area, the basal ganglia (including the ventral striatum), and from basal ganglia through the thalamus to frontal and especially orbitofrontal areas (OFC). Signalling between task-set and motivational assemblies also indicates when various subgoal and final goal states are reached. These actions are constrained by preestablished enduring synaptic connection strengths.

3.4.7. Task-Set Assemblies

The effects of successive activations of similar-shape object assemblies on task performance will differ depending on the task-set assemblies engaged. Consider what might occur in a different kind of task: In some recognition tasks, if a round object was used as a foil for a similarly-shaped object, the time required to reject a similar foil is longer than the time required to reject a different-shape foil (see Engelkamp & Zimmer, 1994, for a review). In that case, activation of the overlapping shape assembly constituents would actually interfere with recognition task performance. In contrast, FE and FR tasks involve the production of a series of responses, and overlap of subassemblies between object assemblies appears to have a net facilitative effect on their successive retrieval. The "cost" of activating different-shaped object assemblies in succession appears to be greater, generally speaking, than successive activation of same-shaped object assemblies (see below). Thus, the task-set assemblies engaged by the task determine the degree to which the facilitation of the same-shape object assemblies' activation results in facilitation of task performance. The remainder of this discussion will focus on mechanisms and principles as applied to the standard FE and FR retrieval task-sets. Task sets are implemented primarily in PFC.

Task-set assemblies coordinate and organize each other's and other (non-task-set) assemblies' activations. Activations of non-task-set assemblies in turn influence task-set assembly operation. Most participants will have not performed the standard laboratory version of FE or FR before. The formation of the particular combination of task-set assemblies (or task-set subassemblies) for a given task that is performed for the first time is a creative process whereby preestablished assemblies are gathered together in new combinations and with some modifications to meet present task goals and subgoals. The rapid changes in synaptic connective strengths required to compose an optimal combination of task-set assemblies for a given task are instigated especially by OFC-centered circuits. Through these OFC-centered changes, some task-set assemblies are strengthened and others weakened, especially during the first trial of FR or the initial stages of an FE trial. Consistent with this claim, Savage et al. (2001) have provided evidence that the level of OFC activity on first-trial FR encoding is strongly positively correlated with subsequent level of categorical clustering in first trial retrieval (also see Miotto et al., 2006). OFC may play an important role in new memory formation (Frey & Petrides, 2002) and in category learning (Reber, Stark, & Squire, 1998).

The regulation of episodic assemblies by task-set assemblies is essential for effective performance in FR and FE. For example, in the recall phase of the FR task, the participant must reactivate the study episode and its contents. Also during the recall phase, the FR participant must keep track of the items already recalled during that trial, and the items that remain to be recalled. Similarly, in FE, the participant must keep track of the items already reported and items yet to be reported. In either FE or FR, the participant must form a retrospective episodic assembly for items already recalled and a prospective episodic assembly for items yet to be recalled. The contents of both the retrospective and prospective assemblies are subject to rapid changes as the person progresses through the retrieval phase of a trial. The task-set assemblies that coordinate and organize these episodic assemblies are adapted from assemblies that are normally active in other common everyday tasks where the person must keep track of what they have done (or said) and what they must do (or say) next.

3.5. Principles of Operation

3.5.1. Redintegration

As Hebb (1949) implied, activation of a shape assembly will enhance retrieval of its respective object assemblies, and vice versa. Activation of an object assembly will facilitate the activation of its word assembly, and vice versa. Activation of an episode assembly (e.g., study phase of the FR list) can facilitate activation of the object assemblies within it, and vice versa. These statements refer to one of the oldest, most well-established principles in memory research. Redintegration is constrained mainly by preestablished enduring synaptic connection strengths and the currently-engaged task-set assembly. Sustained activation of an assembly corresponding to a memory search cue stimulus can also aid redintegration. However, sustained activation is not necessary for redintegration to occur. Shape clustering may occur if activating object assembly A activates its shape assembly C, which in turn activates object assembly B which overlaps on shape assembly C.

3.5.2. Cross-Category Inhibition

When the neurons of shape assembly A fire, the firing of many other different shape assembly neurons will be relatively suppressed to the extent that they lack constituent neurons in shape assembly A. Conversely, when the shape assembly B's neurons increase firing, their increased excitatory activity corresponds with suppression of assembly A, and so on. Local inhibitory connections between neurons and columns of neurons with differing shape-selectivities are likely partly responsible for cross-category inhibition, as suggested by a variety of direct and indirect evidence (Allison, Puce, & McCarthy, 2002; Ferster & Miller, 2000; Kawashima, O'Sullivan, & Roland, 1995; Pelphrey, Mack, Song, Guzeldere, & McCarthy, 2003; Rao, Williams, & Goldman-Rakic, 1999, 2000; Wang, Fujita, & Murayama, 2000). Cross-category inhibition applies at multiple category or domain levels. It may be enhanced by cortico-thalamic-cortico loops, which incorporate additional local thalamic lateral inhibitory effects (LaBerge, 2000). Cross-category inhibition is relative, localized, and temporary: the inhibited neurons do not all fully stop their activity; assemblies linked to the suppressed assembly may not be as suppressed; and some net amount of synaptic enhancement (increased potential for activation) may be retained for the presently suppressed assembly relative to those that are not relevant to a given task. Cross-category inhibition also generally works under the influence of attentional mechanisms. For example, when an FR participant recalls a category, activation of the other to-be-recalled but currently unattended categories will be momentarily relatively suppressed. In fact, this was demonstrated in recall by Polyn et al. (2005), who showed that the level of correspondence between the currently retrieved category's pattern of brain activity and that category's pattern of brain activity at encoding is increased, while there was a reduction in the levels of correspondence between the encoding and retrieval patterns of brain activity for the categories not currently retrieved.

3.5.3. Combinatorial Activation

Individual neurons throughout the brain act as selective gates that regulate the flow of activity within and between assemblies depending on the combination of excitatory and inhibitory synaptic inputs at any given moment. A single cortical neuron may receive thousands to tens of thousands of synaptic inputs and can, in turn, send branching projections with hundreds of terminals onto numerous other neurons and their dendrites. Short-term synaptic plasticity and motivation-related effects on synaptic transmission allow the combinations to vary from moment to moment. However, due to preestablished enduring synaptic connection strengths, there are relatively stable cores of input combinations that fairly reliably activate or deactivate neurons in varying degrees (firing rates), thereby limiting the range of short-term synaptic plasticity dynamics.

Sustained activity serves to keep certain combinatorial gates reliably open or closed for short periods of time to enhance task performance. "Combinatorial activation" means that activation is conditional upon the combination of inputs received. Although most neural theories adopt some notion of combinatorial activation, Damasio's (1989) conception is closest to that of the present theory—though I disagree with his choice of the word "amodal" to describe the format of the codes that bind combinations of stimuli.

3.5.4. Family Resemblance

All else being equal, as the overlap of neurons between assemblies *increases*, the amount of modification or neural processing required for those assemblies to be activated in immediate succession during retrieval *decreases*. As understood here, the degree of family resemblance (cf. Rosch & Mervis, 1975) refers to the amount of overlap of neural instantiations of the representations of discrete features, continuous dimensions, configurations, and sequences. This occurs at the neural population level because of combinatorial activation constraints at the single constituent neuron level; in turn, population-level emergent dynamics can influence combinatorial activation (see Momentum in Successive Retrievals, below) at the level of the individual neuron. Neurally instantiated family resemblance underlying shape representation has been shown dramatically in direct optical imaging of local (microscopic) activity in primate IT (Tsunoda, Yamane, Nishizaki, & Tanifuji, 2001; Wang et al., 1996).

3.5.5. Inertial Activation

The principle of inertial activation operates such that once a particular assembly has been activated, changing the neural assembly's level of activation up or down from the present level is generally more costly neurophysiologically than leaving it at the present level. For elevated inertial levels of activity, this principle is time-limited due to neural fatigue, which occurs at different rates for different neurons, and competing activations. According to the principle of cross-category inhibition, when shape A assembly is active, shape B assembly will be relatively suppressed. To switch from an object assembly-shape A to an object assembly-shape B requires that the excitatory activity of shape A be reduced and that of shape B increased, while the inhibitory input from shape A to shape B be decreased, and shape B's inhibitory input to shape A be increased. All else being equal, then, making the transition in retrieval between object assemblies overlapping considerably on shape assembly A should be less costly and quicker than switching to an object assembly-shape B, because less change is required. In other words, the same or similar shape assembly should remain active unless affected by other factors. This "inertia"-like tendency will be constrained by preestablished enduring synaptic connection strengths, motivation-related effects on synaptic transmission, redintegration, cross-category inhibition, combinatorial activation, and family resemblance. Strong preestablished synaptic connection tendencies, once activated, will be difficult to overcome without some counteracting tendency.

3.5.6. Momentum in Successive Retrievals

Activation momentum is carried to the extent that modifications of the shape assemblies of successively-retrieved object assemblies are not large enough to cause a major change in shape (e.g., as in switching from a sphere to a long rectangular block; though what constitutes a major change in shape is relative to the task context). In retrieval, the transitions of successive shape assembly activations that require only relatively minor modifications can be achieved partly because there is some flexibility in the shape-selectivities of individual neurons (Tanaka, 2002). In addition, the capacity for movement of the local (microscopic) foci of activation across multiple overlapping microcolumns whose neurons have overlapping shape selectivities (as observed by Wang et al., 1996) would also facilitate minor changes in shape from one object assembly activation to the next. These transitions would occur at thousands of local shape-selective sites distributed predominantly in shape-selective parietal and inferior temporal areas. Momentum-in-successive-retrievals will be reduced or disrupted when there is a drastic change in shape from one object assembly activation to the next, due to the greater cost in making such switches (due to inertial activation tendency). Also note that the shape assembly momentum can be reduced by cross-category competition, and neural fatigue (in the currently-active shape assembly); and can be disrupted by signalling between motivational and task-set assemblies.

Neuroimaging and electrophysiological studies indicate that PFC plays a role in regulating activations in PNC (e.g., see Dickerson et al., 2007; Sarnthein, Petsche, Rappelsberger, Shaw, & von Stein, 1998; Tallon-Baudry, Kreiter, & Bertrand 1999). Theoretically, PFC shape-selective sustained activity should be able to influence shape-selective parietal and temporal neurons through the direct long white-fibre cortico-cortico connections between those regions. The PFC sustained activity can thus contribute to the carrying of retrieval momentum for successive object assemblies that have a common or overlapping shape assembly. Moreover, PFC anticipatory/preparatory activity can activate same-shape object assemblies prospectively. This task-set assembly preactivation of the same-shape object assemblies will generate momentum for the to-be-retrieved same-shape object assemblies. Because the PFC sustained activation is distracter-resistant, its engagement should also lead to a pattern of longer clusters in output, whereas sustained activity only in the parietal and temporal areas—activity which is susceptible to disruption—should produce a pattern of shorter clusters.

Larger assemblies with greater family resemblance between them will carry more activation momentum into one another than smaller assemblies with lesser family resemblance. Due to inertia, sustained activity, synaptic enhancement, and other factors, it should be increasingly difficult to counteract this momentum as the overlap between their assemblies and their constituent combinations of subassemblies increases. This partly explains why semantic clustering tends to be much stronger than single-perceptual feature clustering for concrete object items. That is, single-feature perceptual clustering of object items involves the relatively small-scale carry-over of only one common subassembly, and proportionately more switching of subassemblies must occur for same-category transitions. In contrast, concrete semantic clustering involves a massive carry-over of the combinations of multiple subassemblies (including perceptual and motoric), and relatively less switching of subassemblies must occur for same-category transitions.

A host of factors affecting the in-the-moment activations of object assemblies can disrupt within-category momentum. These factors include cross-category inhibition, item distinctiveness effects, serial position effects (in FR), recency of experience with the object or object name prior to the experiment, object word frequency, various nuisance associations and alternative schemes, a slight momentary lapse in the participant's concentration, and so on. The greater the total assembly overlap (i.e., family resemblance) between category members, the more robust the momentum in successive item retrievals within a category in relation to all of those competing factors.

3.5.7. Thresholds and Assembly Locking

In FE and FR tasks, word assemblies' activations need to register a distinctive and memorable trace within the task episode representation and trigger correct word report, but, beyond that, there is little need for extensive activation of the word assemblies. The term "locking" here refers to the dynamic momentary binding between assemblies that is achieved when a task-appropriate critical threshold of an assembly's activation is reached. For example, in the FR task, when the recalled mental image of a volleyball triggers the activation of the word volleyball, then the task-appropriate threshold of activation has been reached. The activation threshold can be measured in terms of the duration, intensity (firing rate), spread (spatial distribution), coherence (of intra-assembly activity), the combinatorial configuration of an assembly's activation, and compatibility of inter-assembly activity (e.g., between a word and object assembly in FR and FE tasks). Locking normally involves only the amount of activation needed for task performance; it does not normally involve complete activation of the whole assembly (cf. assembly "ignition," which is posited to involve complete activation of the whole assembly, see Braitenberg, 1978; Pulvermüller, 1999; Wickens, Hyland, & Anson, 1994). Locking between assemblies is combinatorial/conditional, usually involving mutual interactions between assemblies, such that the specific object assembly activates the correct word assembly (or vice versa) through redintegration. By combinatorial or conditional locking, I mean that a combination of subassemblies within an object assembly, for example, must be activated before that object assembly's corresponding word assembly (another critical combination of subassemblies) is activated.

The occurrence of a (non-chance) shape cluster in FR or FE involves some activation of the object assembly's shape assembly. Significant shape clustering involves locking between the similar (i.e., same-category) shape assemblies. The locking may occur at relatively low levels of (non-conscious, subliminal) activation, or at relatively high levels that reach conscious threshold activation. Regarding momentum in successive retrievals, locking at increased levels of activation should transfer increased activation momentum between assemblies, thereby contributing to stronger clustering. This clustering will be further enhanced by category-selective PFC distracter-resistant sustained activity.

During memory search, the PFC sustained activation of a categorical assembly can evoke the activations of multiple object assemblies at the same time, such that there is the locking of multiple exemplars to the category. These multiple exemplars can be held active, held in queue, and reported sequentially (as described in the answer to Question 3, below).

Individual word assembly and object assembly locking activations must be brief and must break off quickly in FR and FE retrieval. Initiation of an assembly activation, the locking of that assembly with another, and the termination of the assembly activations, are each steps that would be accompanied by phasic firing. The tendency for brief duration, rapid termination of individual word assemblies and object assemblies is regulated by task-set assemblies and motivational assemblies. Termination of the activation of an object or word assembly can be achieved through phasic signalling of the completion of a given subgoal, as regulated by task-set assemblies and motivational assemblies (see below). Activation termination is also needed when making the transition from one clustered category to the next.

3.6. Basic Operations: Clustering and Switching

I offer here a speculative, but plausible, step-by-step account of how shape clustering and switching occur. For simplicity, I will temporarily restrict the description primarily to the frontal cortical and subcortical circuits. It should be kept in mind that the posterior cortical contributions to shape representation, and the MTL contributions to episodic memory, are drawn upon in this processing.

The medial orbitofrontal cortex (mOFC) is critically important in implementing rapid adaptation in responding to changing stimulus-response-reward contingencies, including when stimuli and rewards are of a cognitive or abstract sort. It is not clear whether these neurons have strong preestablished selectivities for shape, but they can take on task-relevant stimulus selectivities in representing specific stimulus-reward contingencies. Lateral PFC (lPFC) areas contain some neurons that show the sustained activity properties mentioned earlier. As the reward value association with respect to the presently-active shape assembly activation changes, mOFC neurons signal the basal ganglia with phasic firing. Medial OFC phasic firing onto the direct pathway (for details, see DeLong, 2000; Joel & Weiner, 2000) should occur when a high reward valence is associated with the presently-activated shape assembly. This input ultimately results in a release of basal ganglial inhibition onto excitatory thalamic neurons, which project back to mOFC, thereby releasing the thalamic excitatory action onto those regions. When a diminishing reward is associated with a shape assembly, phasic input signals the indirect pathway, which further increases the inhibition onto the thalamic excitatory neurons, thereby diminishing their net excitatory input to mOFC. The "winning" shape assembly neurons in mOFC signal lPFC neurons, which implement sustained excitatory activity for the winning shape assemblies. The winning shape assembly should activate its object assemblies by redintegration. Note that, though this discussion focusses on frontal circuits, the same direct/indirect basal ganglial pathway principles apply with respect to its loops with PNC. However, PNC neurons' sustained activity does not seem to be distracter-resistant like that of lPFC (Desimone, 1996).

Given cross-category inhibition effects between directly connected local cortical areas, the presently-active shape assembly associated with high reward will suppress the activation of the alternative shape assemblies. The presently-active high reward shape assembly neurons, and the presently-less-active low reward shape assembly neurons, can all have their present (inertial) level of activity sustained temporarily by lPFC neurons. The lPFC neurons can also respond selectively for stimulus-response-reward contingencies, but the representation of rapid changes in these contingencies appears to be more dependent on mOFC. As the high-reward shape assembly neurons instigate their successive object assembly and word assembly lockings, words are reported in clusters as long as this inertial tendency continues to win out over other-shape-category and other competitive activation influences. When the reward value of the presently-active shape assembly begins to drop substantially (as signaled by performance monitoring feedback, probably from anterior cingulate (AC) areas), mOFC neurons phasically signal the indirect pathway, which reduces the thalamic excitatory feedback to the presently most active shape assembly neurons. This reduction is accompanied by an increase in activation for the remaining to-be-retrieved shape assemblies, with the winner of the competitive activation inhibiting the other shape assemblies and signaling lPFC sustained activity for the winning shape assembly. This momentum shift corresponds with switching between categories. The process continues in this fashion for the successive shape assembly activations.

3.7. First Answers to Difficult Questions

1. How is reward value registered?

Reward value is determined partly by the goals and subgoals of the task, whose representations are implemented in the task set assemblies. Performance monitoring of retrieval activations and of word articulations is centered in medial frontal cortex, including the AC. One possibility is that as retrieval and word report delays increase, AC signals mOFC neurons that represent decreases in reward value associated with a shape assembly's activation. The change evokes phasic firing in mOFC, which signals lPFC and the basal ganglia. Note that the mOFC receives input from the subcortical affective/motivational/reward-related structures, including the amygdala, and the ventral tegmental area, which connects to nucleus accumbens, which in turn projects to dorsomedial (DM) thalamus and then to lPFC/mOFC. Reward-stimulus category contingencies are registered in episode assemblies, which are centered in areas such as hippocampus, amygdala, and surrounding medial temporal cortex, and which receive input from and project to mOFC. The DM thalamus also receives input from the amygdala. These stimulus-reward contingent neurons (especially in mOFC) can be reactivated from the episode assemblies later in a trial or from the more generalized (and task-set-organized) episode assembly that develops in trials of retrieval after the first trial. As the number of trials increases, the episode assembly becomes increasingly script-like (i.e., routine, well-rehearsed).

2. How do the cortical-basal ganglial-thalamic-cortico loop interactions maintain the specificity of the operations for different shape assemblies?

First, neurons with similar selectivities tend to "fire and wire" together; they exhibit selectivities precisely because of the combination of inputs they receive (i.e., they are constrained by combinatorial activation). Thus, the shape-reward mOFC neurons and the shape-selective neurons in lPFC should be connected. (This is in addition to the posterior cortical shape assembly neurons' inputs to lPFC/mOFC). Second, there are direct bidirectional connections between the parvocellular DM thalamus and lPFC neurons, and between magnocellular DM thalamus and mOFC neurons. Those frontal cortical projections, given their response-reward-stimulus-selectivities, can effectively bias different DM thalamic neurons' activity levels. This additional action onto the DM thalamus neurons will further enhance the specificity of the excitatory and suppressive shape assembly actions in lPFC and mOFC. In effect, the DM thalamus works as a combinatorial filter that is regulated by lPFC and mOFC activity which, in turn, is regulated by DM thalamic activity.

3. How are the object assembly and word assembly activations ordered into the linear sequence for response output?

The above-described direct/indirect pathway architecture can implement the sequential activations of word assemblies in much the same way as successive clusters and switches are implemented. In shape clustering, there is no restriction on the exact ordering of the individual words within shape clusters. Nevertheless, some kind of flexible task-set assembly is needed to assign the objects and words to response sequence positions. An old, but in my opinion, still valid idea is that the response sequence is organized into spatio-temporal "slots" (Lashley, 1951; Luria, 1970). Neural selectivities operate according to combinatorial activation, and support sequential task-set assembly components. These task element and task element combination neurons are based to a considerable extent in higher medial frontal motor (MFM) areas and parietal cortex, and have been demonstrated in macaques (for a review, see Tanji, 2001); and the same areas are activated when humans do sequential-order tasks. Neurons that are selective for the *rank order* of an action in a sequence or prospective sequence, but non-selective for the specific action itself, have also been found in macaques (Tanji, 2001). In addition, neurons that are selective for remembering the sequential order in which objects are presented have been found in the lPFC in macaques (Ninokura, Mushiake, & Tanji, 2003). In that study, the monkey had to remember the order in which a yellow circle, blue rectangle, and red cross were presented, and then, after a brief delay, had to touch the objects in that same order.

The higher MFM, in addition to Broca's area (which is in lPFC), is important in speech planning and production in humans. Object and word representations could be assigned to sequentially-ordered slots, with the lPFC neurons representing the conjunction of word and serial position. These conjunctive representations of word (partly implemented in Broca's area neurons) and output position (in MFM) could be assigned prospectively, in preparation for output, and held in queue briefly and prospectively by lPFC sustained neurons. The above-described basal ganglial direct/indirect pathway operations could help instigate successive word assembly activations and deactivations in reporting successive object names according to their positions in the ordering of spatiotemporal slots in the task set representation. The order in which categories or clusters are reported in sequence could also be implemented through the assignment of categories to spatiotemporal slots in the task set assembly formed by the participant within the FE or FR task context.

4. How does shape clustering differ from spatial relational and serial clustering?

Shape categorical clustering differs from spatial and serial clustering of object nouns with respect to the (a) underlying neural assemblies that organize the information, and (b) the restrictions these assemblies put on the sequential order of word outputs. Spatial context assemblies are organized according to spatial relationships between components as specified through learning, but there should not be very strong unidirectionally-ordered sequence biases, independently of sequential action relations between the component elements. Note that in spatial contextual clustering, as discussed here, objects (object assemblies) are the elements of the arrangement in the spatial context. These biases should be understood in terms of combinatorial activation, whereby neurons are conditionally-selective for the relation between items. Thus, when the sets of assemblies needed for sequential output become engaged, there is already a preestablished bias for the binding of object assemblies (and word assemblies) to prospective rank order of output. Activation of spatial contextual assemblies constrains the organization of object assemblies to rank order outputs. (And then the object assemblies' order assignments determine those of their corresponding word assemblies). This also makes sense in terms of perceptual theoretic assumptions because there is a reactivation of the assemblies that tended to be reliably engaged across the original experiences (Barsalou, 1999), and this reactivation is in preparation for actions in contexts (Glenberg, 1997). Activating objects in the order in which they were–or are likely to be–encountered is more useful than would be the activation of a randomly-ordered list of items. For example, consider writing up a prospective grocery list before going to your familiar supermarket. One strategy would be to list the items according to their positions in the supermarket, positions where you expect to encounter them again. (There are, of course, many other factors that will affect list organization).

Serial clustering in FR is based on the reactivation of interitem temporal contiguities as organized by the list study or prior retrieval trial episode assemblies and by any other serial contingencies between items that were established prior to the experimental episode. These episode assembly links, whose formation is centred in the MTL, bias the assignment of object assemblies (and hence word assemblies) to rank order output positions such that there are stronger restrictions of temporal directional order of output, relative to the restrictions specified by spatial context assemblies and overlapping shape assemblies. Serial organization should be implemented, in part, by sequence-selective neurons in MFM (e.g., neurons that are selective such that they will respond maximally for sequence A-B-C, but not for C-B-A, A-C-B, etc., see Tanji, 2001). To return to the grocery store example, clusters of items may be listed by the sequential order in which those groups of items tend to be encountered in one's usual route through the store (e.g., meats, then dairy, then produce, then cereals...). Within those groups, items may be recalled in the order in which they tend to be picked up.

5. Is overlap of shape assemblies the only way similarly-shaped items can be clustered?

First, shape clustering can also be implemented with minimal actual overlap of shape-selective neurons through associative linking of items by MTL neurons. Note that MTL neurons can develop selectivities for pairs of stimuli (for a review, see Miyashita, 1993). Possibly, selective attention to any salient similarity, contrast, or relation between items could induce an association between a pair of items. Reactivation of overlapping shape-selective neurons could activate familiarity- or recognition-based neurons in the MTL, thus linking same-shape items across long interrupted temporal delays during list encoding. These encoded "microepisodes" could then be reactivated with the task list episode assembly during retrieval, thereby facilitating clustering of the same-shape items. It is also conceivable that same-shape items could be linked only via some common association, such as a shape name (word assembly) or shared context (spatial contextual assembly), and that these associations could produce apparent shape clustering. (This latter possibility is similar to the apparent shape clustering produced by nuisance associations within shape categories).

Second, with minimal overlap, shape clustering can be implemented through shape selectivities in lPFC and lateral OFC that may be more criterion- than overlap-based. Thus, if a given shape fits a neuron's shape categorical criterion for activation, sustained shape-selective lPFC activity can hold a subset of shape-selective PNC neurons active (see Momentum in Successive Retrievals). Note that, in monkeys trained to classify complex three-dimensional forms (cat versus dog forms), many PFC shape-selective neurons can show wide tolerance for shape variation within category but also show abrupt offset of activation for shapes just outside the criterial boundaries (Freedman et al., 2001). I believe that normally PFC, MTL, and PNC all contribute to clustering.

Third, there may be separate patches of neurons that are selective for approximately the same shape but which are located in different parts of the temporal and parietal areas. If the activation of one of those patches selective for, say, spheroids, triggers sustained activity in shape-selective neurons in PFC, the sustained activity in PFC could in turn trigger the activation of other spheroid-selective patches of neurons in other temporal and parietal areas. Note that these interactions would be made through long white-fibre connections between PFC and temporal and parietal areas.

6. How does the present theory account for the clustering of "opposites"?

I suspect that if the shape words *square* and *circle* were included in a study list containing mostly non-shape words, the shape words would cluster in retrieval more than would be expected by chance. The effect could be attributed partly to word associations, but a more complete explanation is needed that takes into account cross-category inhibition. A circle-shape assembly and a square-shape assembly are themselves members of the shape domain, a larger and more inclusive assembly wherein different shapes overlap some of the same neural constituents. Square and circle differ with respect to curvature and other aspects, but are aligned (see Medin et al., 1993) with respect to shape, and are more similar to each other than to items in other modalities (e.g., colours, sounds, etc.). Thus, the cross-category inhibitory effects are relative to the level of the contrast set activated in the task. The principle of cross-category inhibition in operation in the FR of a list operates on at least two levels: At one level, cross-category inhibition contributes to contrast between distinctive exemplars; at the next level, cross-category inhibition contributes to contrast between categories. (The terms "categories" and "exemplars" are relative to the stimuli used and the task-set assemblies engaged). For a different FR list containing shape names such as circle, square, and triangle, within a list of other-modality category names such as for colours, sounds, and tastes, items should cluster in retrieval according to their respective modalities.

3.8. Summary

Retrieving an object from LTM entails reactivating a subset of the neurons that were active in previous experience with the object, including neurons selective for perceptual properties such as the object's shape. It is argued here that the clustering of objects by shape similarity in FE and FR retrieval occurs mainly due to the overlap between object assemblies that share parts of the same shape assembly. Generally, the greater the overlap between neural assemblies, the more likely that the activation of one will lead to activation of the other, in FE and FR. Indeed, simply by activating an object assembly and its shape assembly, part of the work is already accomplished in activating another object assembly that overlaps the shape assembly of the former. Shape-selective sustained activity in PNC and PFC—the latter of which can continue without disruption from the activations of intervening items—is suggested as one of the major contributors to shape clustering. Shape-selective sustained activity prolongs the activation of a shape category assembly and thus increases the chances of triggering the successive activations of object assemblies that overlap that shape assembly. In addition, once a shape assembly has been activated, there may be short-term synaptic enhancement (STSE) for the neurons in that shape assembly, increasing the likelihood of its reactivation and, hence, activation of the object assemblies that overlap it. The relative numbers and strengths of preestablished enduring synaptic connections will partly determine which organizational scheme(s) (e.g., shape similarity, spatial contextual, semantic categorical, etc.) participants may use effectively to enhance performance at any given stage in the task episode. Task-set assemblies assign words, objects, and shape categories to spatio-temporal slots, permitting items and clusters to be reported in a linear sequence.

References for Chapter 3

Allison, T., Puce, A, & McCarthy, G. (2002). Category-sensitive excitatory and inhibitory processes in human extrastriate cortex. *Journal of Neurophysiology, 88,* RC 2864-2868.

Andersen, R. A., Bracewell, R. M., Barash, S., Gnadt, J. W., & Fogassi, L. (1990). Eye position effects on visual, memory, and saccade-related activity in areas LIP and 7a of macaque. *Journal of Neuroscience, 10,* 1176-1196.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22,* 577-660.

Becker, S., & Lim, J. (2003). A computational model of prefrontal control in free recall: Strategic memory use in the California Verbal Learning Task. *Journal of Cognitive Neuroscience, 15,* 821-832.

Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *The Journal of General Psychology, 49,* 229-240

Bousfield, W. A., & Cohen, B. H. (1953). The effects of reinforcement on the occurrence of clustering in the recall of randomly arranged associates. *The Journal of Psychology, 36,* 67-81.

Bodner, M., Kroger, J., & Fuster, J. M. (1996). Auditory memory cells in dorsolateral prefrontal cortex. *NeuroReport, 7,* 1905-1908.

Braitenberg, V. (1978). Cell assemblies in the cerebral cortex. In R. Heim, & G. Palm, (Eds.), *Lecture notes in biomathematics: 21. Theoretical approaches to complex systems* (pp. 171-188). New York: Springer-Verlag.

Courtney, S. M., Petit, L., Maisog, J. M., Ungerleider, L. G., & Haxby, J. V. (1998). An area specialized for spatial working memory in human frontal cortex. *Science, 279,* 1347-1351.

Damasio, A. R. (1989). Time-locked multiregional retroactivation: A systems-level proposal for the neural substrate of recall and recognition. *Cognition, 33,* 25-62.

DeLong, M. R. (2000). The basal ganglia. In E. R. Kandel, J. H. Schwartz, & T. H. Jessell, (Eds.), *Principles of neural science* (4th ed.). (pp. 853-867). New York: McGraw-Hill.

Desimone, R. (1996). Neural mechanisms for visual memory and their role in attention. *Proceedings of the National Academy of Sciences, USA, 93,* 13494-13499.

Dickerson, B. C., Miller, S. L., Greve, D. N., Dale, A.M., Alvert, M. S., Schacter, D. L., et al., (2007). Prefrontal-hippocampal-fusiform activity during encoding predicts intraindividual differences in free recall ability: An event-related functional-anatomic MRI study. *Hippocampus, 17,* 1060-1070.

Egorov, A. V., Hamam, B. N., Fransen, E., Hasselmo, M. E., & Alonso, A. A. (2002). Graded persistent activity in entorhinal cortex neurons. *Nature, 420,* 173-178.

Ehrsson, H. H., Geyer, S., & Naito, E. (2003). Imagery of voluntary movement of fingers, toes, and tongue activates corresponding body-part-specific motor representations. *Journal of Neurophysiology, 90,* 3304-3316.

Engelkamp, J., & Zimmer, H. D. (1994). *Human memory: A multimodal approach.* Toronto: Hogrefe and Huber.

Farah, M. J. (1990). *Visual agnosia: Disorders of object recognition and what they tell us about normal vision.* Cambridge, MA: MIT Press.

Fellows, L. K. & Farah, M. J. (2003). Ventromedial frontal cortex mediates affective shifting in humans: Evidence from a reversal paradigm. *Brain, 126,* 1830-1837.

Ferster, D., & Miller, K. D. (2000). Neural mechanisms of orientation selectivity in the visual cortex. *Annual Review of Neuroscience, 23,* 441-471.

Fisher, S. A., Fischer, T. M., & Carew, T. J. (1997). Multiple overlapping processes underlying short-term synaptic enhancement. *Trends in Neurosciences, 20,* 170-177.

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. *Science, 291,* 312-316.

Frey, S. & Petrides, M. (2002). Orbitofrontal cortex and memory formation. *Neuron, 36,* 171-176.

Funahashi, S., Bruce, C. J., & Goldman-Rakic, P. S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology, 61,* 331-349.

Fuster, J. (1995). *Memory in the cerebral cortex: An empirical approach to neural networks in the human and nonhuman primate.* Cambridge, MA: Bradford, MIT Press.

Fuster, J. M., & Jervey, J. P. (1981). Inferotemporal neurons distinguish and retain behaviorally relevant features of visual stimuli. *Science, 212,* 952-955.

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences, 20,* 1-55.

Handy, T. C., Miller, M. B., Schott, B., Shroff, N. M., Janata, P., Van Horn, J. D., et al., (2004). Visual imagery and memory: Do retrieval strategies affect what the mind's eye sees? *European Journal of Cognitive Psychology, 16,* 631-652.

Hauk, O, Johnsrude, I. S., & Pulvermüller, F. (2004) Somatotopic representation of action words in human motor and premotor cortex. *Neuron, 41,* 301-7.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science, 293,* 2425-2430.

Hebb, D. O. (1949). *The organization of behavior.* New York: John Wiley & Sons.

Hebb, D. O. (1968). Concerning imagery. *Psychological Review, 75,* 466-477.

Heil, M., Rösler, F., & Hennighausen, E. (1996). Topographically distinct cortical activation in episodic long-term memory: The retrieval of spatial versus verbal information. *Memory & Cognition, 24,* 777–795.

Humphreys, G. W., & Forde, E. M. E. (2001). Hierarchies, similarity, and interactivity in object recognition: "Category-specific" neuropsychological deficits. *Behavioral and Brain Sciences, 24,* 453-509.

Huttenlocher, P. R. (2002). *Neural plasticity: The effects of the environment on the development of the cerebral cortex.* Cambridge, MA: Harvard University Press.

Ishai, A., Ungerleider, & Haxby, J. V. (2000). Distributed neural systems for the generation of visual images. *Neuron, 28,* 979-990.

Ishai, A., Ungerleider, L. G., Martin, A., Schouten, J. L., & Haxby, J. V. (1999). Distributed representation of objects in the human ventral visual pathway. *Proceedings of the National Academy of Sciences, USA, 96,* 9379-9384.

Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Brain and Behavioral Sciences, 17,* 187-245.

Joel, D., & Weiner, I. (2000). Striatal contention scheduling and the split-circuit scheme of basal ganglial-thalamocortical circuitry: From anatomy to behavior. In R. Miller and J. R. Wickens, (Eds.), *Brain dynamics and the striatal complex* (pp. 209-236).

Kawashima, R., O'Sullivan, B. T., & Roland, P. E. (1995). Positron emission tomographic studies of cross-modality inhibition in selective attentional tasks: Closing the "mind's eye." *Proceedings of the National Academy of Sciences, USA, 92,* 5969-5972.

Kosslyn, S. M. (1994). *Image and brain: The resolution of the imagery debate.* Cambridge, MA: Bradford, MIT Press.

Kreiman, G., Koch, C., & Fried, I. (2000a). Category-specific visual responses of single neurons in the human medial temporal lobe. *Nature Neuroscience, 3,* 946-953.

Kreiman, G., Koch, C., & Fried, I. (2000b). Imagery neurons in the human brain. *Nature, 408,* 357-361.

LaBerge, D. (2000). Networks of attention. In M. S. Gazzaniga, (Ed.), *The new cognitive neurosciences* (2nd ed., pp. 711-724). Cambridge, MA: Bradford, MIT Press.

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress, (Ed.), *Cerebral mechanisms in behavior: The Hixon symposium* (pp. 112-136). New York: Wiley.

Leung, H.-C., Gore, J. C., & Goldman-Rakic, P. S. (2002). Sustained mnemonic response in the human middle frontal gyrus during on-line storage of spatial memoranda. *Journal of Cognitive Neuroscience, 14,* 659-671.

Levine, D. N., Warach, J., and Farah, M. J. (1985). Two visual systems in mental imagery: Dissociation of "What" and "Where" in imagery disorders due to bilateral posterior cerebral lesions. *Neurology, 35,* 1010-1018.

Levy, R., Friedman, H. R., Davachi, L., & Goldman-Rakic, P. S. (1997). Differential activation of the caudate nucleus in primates performing spatial and nonspatial working memory tasks. *The Journal of Neuroscience, 17,* 3870-3882.

Luria, A. R. (1970, March). The functional organization of the brain. *Scientific American,* 66-73.

Martin, A. (2001). Functional neuroimaging of semantic memory. In R. Cabeza, & A. Kingstone, (Eds.), *Handbook of functional neuroimaging of cognition* (pp. 153-186). Cambridge, MA: Bradford, MIT Press.

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology, 58,* 25-45.

Martin, A., Ungerleider, L. G., & Haxby, J. V. (2000). Category specificity and the brain: The sensory/motor model of semantic representations of objects. In M.S. Gazzaniga, (Ed.), *The new cognitive neurosciences* (2nd ed., pp. 1023-1037). Cambridge, MA: Bradford, MIT Press.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100*, 254-278.

Miller, R. & Wickens, J. R. (Eds.). (2000). *Brain dynamics and the striatal complex.* Amsterdam, Netherlands: Harwood Academic Publishers.

Milner, P. M. (1999). *The autonomous brain: A neural theory of attention and learning.* Mahwah, NJ: Lawrence Erlbaum Associates.

Miotto, E. C., Savage, C. R., Evans, J. J., Wilson, B. A., Martins, M. G. M., Iaki, S., et al. (2006). Bilateral activation of the prefrontal cortex after strategic semantic cognitive training. *Human Brain Mapping 27*, 288-295.

Miyashita, Y. (1993). Inferior temporal cortex: Where perception meets memory. *Annual Review of Neuroscience, 16*, 245-263.

Moscovitch, M., Kapur, S., Kohler, S., & Houle, S. (1995). Distinct neural correlates of visual long-term memory for spatial location and object identity: A positron emission tomography study in humans. *Proceedings of the National Academy of Sciences, USA, 92*, 3721–3725.

Ninokura, Y., Mushiake, H., & Tanji, J. (2003). Representation of the temporal order of visual objects in the primate lateral prefrontal cortex. *Journal of Neurophysiology, 89*, 2868-2873.

Nyberg, L., Habib, R., McIntosh, A. R., & Tulving, E. (2000). Reactivation of encoding-related brain activity during memory retrieval. *Proceedings of the National Academy of Sciences, USA, 97*, 11120-11124.

O'Doherty, J. O., Critchley, H., Deichman, R., & Dolan, R. (2003). Dissociating valence of outcome from behavioral control in human orbitofrontal and ventral prefrontal cortices. *Journal of Neuroscience, 23*, 7931-7939.

Oliver, R. T., and Thompson-Schill, S. L (2003). Dorsal stream activation during retrieval of object size and shape. *Cognitive, Affective, & Behavioral Neuroscience, 3*, 309-322.

Ó Scalaidhe, S. P., Wilson, F. A. W., & Goldman-Rakic, P. S. (1999). Face-selective neurons during passive viewing and working memory performance of rhesus monkeys: Evidence for intrinsic specialization of neuronal coding. *Cerebral Cortex, 9*, 459-475.

Pelphrey, K. A., Mack, P., Song, A., Güzeldere, G., & McCarthy, G. (2003). Faces evoke spatially differentiated patterns of BOLD activation and deactivation. *NeuroReport, 14*, 955-959.

Penfield, W., & Perot, P. (1963). The brain's record of auditory and visual experience. A final summary and discussion. *Brain, 86*, 595–696.

Penfield, W., & Rasmussen, T. (1950). *The cerebral cortex of man: A clinical study of localization.* Boston: Little, Brown, & Company.

Polyn, S. M., Natu, V. S., Cohen, J. D., & Norman, K. A. (2005). Category-specific cortical activity precedes retrieval during memory search. *Science, 310*, 1963-1966.

Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences, 22*, 253-336.

Quintana, J., Yejeya, J., & Fuster, J. M. (1988). Prefrontal representation of stimulus attributes during delay tasks. I. Unit activity in cross-temporal integration of sensory and sensory-motor information. *Brain Research, 474*, 211-221.

Rao, S., Rainer, G., & Miller, E. K. (1997). Integration of what and where in the primate prefrontal cortex. *Science, 276*, 821-824.

Rao, S., Williams, G. V., & Goldman-Rakic, P. S. (1999). Isodirectional tuning of adjacent interneurons and pyramidal cells during working memory: Evidence for microcolumnar organization in PFC. *Journal of Neurophysiology, 81*, 1903-1916.

Rao, S., Williams, G. V., & Goldman-Rakic, P. S. (2000). Destruction and creation of spatial tuning by disinhibition: $GABA_A$ blockade of prefrontal cortical neurons engaged by working memory. *Journal of Neuroscience, 20*, 485-494.

Reber, P. J., Stark, C. E. L., & Squire, L. R. (1998). Cortical areas supporting category learning identified using functional MRI. *Proceedings of the National Academy of Sciences, USA, 95*, 747-750.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences, 21*, 188-194.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2000). Cortical mechanisms subserving object grasping and action recognition: A new view on the cortical motor functions. In M. S. Gazzaniga, (Ed.), *The new cognitive neurosciences* (2nd ed., pp. 539-552). Cambridge, MA: Bradford, MIT Press.

Roland, P. E., & Friberg, L. (1985). Localization of cortical areas activated by thinking. *Journal of Neurophysiology, 53,* 1219–1243.

Rolls, E. T., & Treves, A. (1998). *Neural networks and brain function.* Oxford: Oxford University Press.

Romo, R., Brody, C. D., Hernández, A., & Lemus, L. (1999). Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature, 399,* 470-473.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7,* 573-605.

Rosenzweig, M. R., Leiman, A. L., & Breedlove, S. M. (1999). *Biological psychology: An introduction to behavioral, cognitive, and clinical neuroscience* (2nd ed.). Sunderland, MA: Sinauer Associates, Inc.

Rösler, F., Heil, M., & Hennighausen, E. (1995). Distinct cortical activation patterns during long-term memory retrieval of verbal, spatial and color information. *Journal of Cognitive Neuroscience, 7,* 51–65.

Sakai, K. & Miyashita, Y. (1991). Neural organization for the long-term memory of paired associates. *Nature, 354,* 152-155.

Sakata, H., Taira, M., Kusunoki, M., Murata, A., Tanaka, Y., & Tsutsui, K. (1999). Neural coding of 3D features of objects for hand action in the parietal cortex of the monkey. In G. W. Humphreys, J. Duncan, & A. Treisman, (Eds.), *Attention, space, and action: Studies of cognitive neuroscience.* New York: Oxford University Press.

Sakagami, K. & Watanabe, M. (2007). Integration of cognitive and motivational information in the primate lateral prefrontal cortex. *Annals of the New York Academy of Sciences, 1104,* 89-107.

Samejima, K. & Doya, K. (2007). Multiple representations of belief states and action values in corticobasal ganglia loops. *Annals of the New York Academy of Sciences, 1104,* 213-228.

Sarnthein, J., Petsche, H., Rappelsberger, P., Shaw, G. L., & von Stein, A. (1998). Synchronization between prefrontal and posterior association cortex during human working memory. *Proceedings of the National Academy of Sciences, USA, 95,* 7092-7096.

Savage, C. R., Deckersbach, T., Heckers, S., Wagner, A.D., Schacter, D. L., Alpert, N. M., et al. (2001). Prefrontal regions supporting spontaneous and directed application of verbal learning strategies: Evidence from PET. *Brain, 124,* 219-231.

Schneggenburger, R., Sakaba, T., & Neher, E. (2002). Vesicle pools and short-term synaptic depression: Lessons from a large synapse. *Trends in Neurosciences, 25,* 206-212.

Sereno, A. B., & Maunsell, J. H. R. (1998). Shape selectivity in the primate intraparietal cortex. *Nature, 395,* 500-503.

Simmons, K. W., Martin, A., & Barsalou, L. W. (2005). Pictures of appetizing foods activate gustatory cortices for taste and reward. *Cerebral Cortex, 15,* 1602-1608.

Stent, G. S. (1973). A physiological mechanism for Hebb's postulate of learning. *Proceedings of the National Academy of Sciences, USA, 70,* 997-1001.

Tallon-Baudry, C., Kreiter, A., & Bertrand, O. (1999). Sustained and transient oscillatory responses in the gamma and beta bands in a visual short-term memory task in humans. *Visual Neuroscience, 16,* 449–459.

Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience, 19,* 109-139.

Tanaka, K. (2002). Neuronal representation of object images and effects of learning. In M. Fahle and T. Poggio, (Eds.), *Perceptual learning* (pp. 67-82). Cambridge, MA: Bradford, MIT Press.

Tanji, J. (2001). Sequential organization of multiple movements: Involvement of cortical motor areas. *Annual Review of Neuroscience, 24,* 631-651.

Thompson-Schill, S. L., Aguirre, G. K., D'Esposito, M., & Farah, M. J. (1999). A neural basis for category and modality specificity in semantic knowledge. *Neuropsychologia, 37,* 671–676.

Tsunoda, K., Yamane, Y., Nishizaki, M., & Tanifuji, M. (2001). Complex objects are represented in macaque inferotemporal cortex by the combination of feature columns. *Nature Neuroscience, 4,* 832-838.

Vaidya, C. J., Zhao, M., Desmond, J. E., & Gabrieli, J. D. E. (2002). Evidence for cortical encoding specificity in episodic memory: Memory-induced re-activation of picture processing areas. *Neuropsychologia, 40,* 2136-2143.

Vanderberghe, R., Price, C. J., Wise, R., Josephs, O., Frackowiak, R. S. J., (1996). Functional anatomy of a common semantic system for words and pictures. *Nature, 383,* 254-256.

Wallis, J. D. & Miller, E. K. (2003). Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *European Journal of Neuroscience, 18*, 2069-2081.

Wang, G., Tanaka, K., & Tanifuji, M. (1996). Optical imaging of functional organization in the monkey inferotemporal cortex. *Science, 272*, 1665-1668.

Wang, Y., Fujita, I., & Murayama, Y. (2000). Neuronal mechanisms of selectivity for object features revealed by blocking inhibition in inferotemporal cortex. *Nature Neuroscience, 3*, 807-813.

Warrington, E. K. & McCarthy, R. A. (1987). Categories of knowledge: Further fractionations and an attempted integration. *Brain, 110*, 1273-1296.

Wheeler, M. E., Petersen, S. E., Buckner, R. L. (2000). Memory's echo: Vivid remembering reactivates sensory-specific cortex. *Proceedings of the National Academy of Sciences, USA, 97*, 11125-11129.

Wickens, J., Hyland, B., & Anson, G. (1994). Cortical cell assemblies: A possible mechanism for motor programs. *Journal of Motor Behavior, 26*, 66-82.

Yoo, S-S., Freeman, D. K., McCarthy, J. J. III, & Jolesz, F. A. (2003). Neural substrates of tactile imagery: A functional MRI study. *NeuroReport, 14*, 581-585.

Zatorre, R. J., & Halpern, A. R. (2005). Mental concerts: Musical imagery and auditory cortex. *Neuron, 47*, 9–12.

Zatorre, R.J., Halpern, A. R., Perry, D. W., Meyer, E. & Evans, A. C. (1996). Hearing in the mind's ear: A PET investigation of musical imagery and perception. *Journal of Cognitive Neuroscience, 8*, 29-46.

Appendix A

FE Study, Phase 3: Object Listing and Typicality Rating

Object Listing and Shape Typicality Rating

The object listing and shape typicality rating tasks were used to obtain item dominance and typicality information for object nouns that could be used in the subsequent FR study. Six individually-tested participant-judges (3 males, 3 females) each first completed 17 brief object listing tasks and then a shape typicality rating of a selection of object names that had been compiled from Phase 1 FE output. This part of the FE study (i.e., Phase 3) took place after the completion of the shape classifications of the master sets. A participant's total session took approximately two hours.

Object Listing

The purpose of this task was to obtain from participants some additional object names for shape categories to possibly be used in the FR study. A secondary purpose was to obtain an estimate of how many object names people could produce for the various specific shape categories within a brief period of time.

A stack of 17 pages was used for the object listing, where each page had a description of a specific shape category at the top and 20 lined spaces beneath. The participant's task was to write the names of as many objects as she/he could think of, within 60 seconds, that were examples of the specific shape category. Before the object listing tasks proper were started, the experimenter explained the definition of a "real physical thing," which was the same as that used for the Phase 1 FE practice task. Next, in a familiarization phase, the experimenter went through the entire stack with the participant, and ensured that he/she understood each category description by asking her/him to write the name of one example for that category before moving on to the next category. The object listing task instructions were very similar to those used in the FE task in Phase 1 of the FE study, aside from the more specific definitions of the categories.

Seventeen shape domains were used. These included (1) flat rectangular, (2) slab rectangular, (3) block-like rectangular—not long but longer than cubical, (4) cubical block-like rectangular, (5) long board-like rectangular, (6) long cylinders, (7) long pointed cylinders, (8) cylinders that are not long, (9) cup-like cylinders, (10) spheroid, (11) bulbous, (12) bowl-like, (13) rings, (14) disks, (15) cones, (16) string-like, and (17) strip-like. Similar domains (e.g., rectangular) were grouped together in the sequence of pages. Otherwise, the order of the 17 domains was random. Definitions for long and flat were consistent with those for the original superordinate FE domains. The 17 domains were selected according to distinctions made by the master list shape judges in their classifications and descriptions, though cones was added by the researcher based on previous pilot research involving object listing. The sample captured the major subcategories that could possibly contain enough items to be used subsequently in composing the FR lists.

Each of the 17 trials began with the experimenter starting the stopwatch as soon as the participant had read the instructions at the top of the page. The experimenter stopped each trial when 60 s had elapsed. There was a 15 s break between trials.

Object Listing Results

The mean numbers of objects listed for each shape category and for each participant were calculated. The first example object name generated by the participant during the preliminary familiarization phase was excluded from the calculations. The mean numbers of items produced per category within 60 seconds are as follows (M, *SD*): flat rectangular (4.0, *2.45*), slab rectangular (2.3, *1.51*), block-like rectangular (4.3, *2.07*), cubical block-like rectangular (4.2, *2.64*), long board-like rectangular (3.0, *3.22*), long cylinders (4.0, *2.28*), long pointed cylinders (3.7, *1.97*), cylinders that are not long (2.7, *3.14*), cup-like cylinders (2.5, *1.76*), spheroid (6.5, *2.17*), bulbous (2.2, *1.60*), bowl-like (4.3, *1.97*), rings (3.7, *1.75*), disks (3.3, *2.07*), cones (2.5, *1.76*) , string-like (5.8, *1.72*) , and strip-like (3.7, *2.94*). Clearly, more objects were listed for spheroid and string-like than for other categories.

The mean numbers of items produced per category by the 6 participants are as follows (mean, *SD*): #1 (6.9, *1.96*), #2 (3.4, *2.12*), #3 (2.1, *1.43*), #4 (3.4, *1.97*), #5 (3.4, *1.37*), #6 (3.1, *2.11*). Overall, participants produced a mean of 3.7 items per category within 60 seconds. Out of 102 trials (6 participants x 17 categories), there were four trials that yielded zero (0) items, and 17 trials that yielded only one item. This finding, together with the overall finding of low production, suggests that, in subsequent research, modifications to various aspects of the task may be needed to increase production. Due to low production, the researcher did not deem it appropriate to use these results for estimating the category dominance of the items.

Shape Typicality Rating

For each participant, a stack of several pages was prepared containing approximately 350 different object names in total, organized into 17 shape categories specified from the previous (different) judges' shape sorting tasks. The sample of objects was drawn almost entirely from the Phase 1 FE output. Most of these object names had been selected according to the FR list criteria, but, to provide some variety, a smaller number of them were also included that were likely to be considered poorer examples of their respective categories. Some items were also added from previous pilot work. Similar domains (e.g., rectangular) were kept together in the sequence of pages. Otherwise, the order of domains in the stack was randomized. Three different orderings of the item set were prepared, each containing different randomizations of the individual item positions within category lists.

The participant had already been familiarized with the categories from the object listing tasks. The task was simply to rate, by circling the appropriate number on a scale, how closely each item on the list conformed to her/his ideal notion of the shape category in question. The scale ranged from 1 to 7, where 1 indicated very low conformity, and 7 indicated very high conformity to the ideal. As with the similarity rating task, the experimenter checked over the participant's ratings, made suggestions where appropriate, but the participant decided on the final ratings.

Shape Typicality Instructions

The shape typicality rating instructions were as follows:

> On these pages are a large number of the names of objects that are grouped by shape category. There is a scale ranging from 1 to 7. A rating of 7 for an object indicates that it conforms very highly with the idealized geometric form of the category, whereas a rating of 1 indicates that there is very low conformity to that ideal. The degree of conformity of an object's shape to the category ideal will be rated as follows: 1 = very low; 2 = low; 3 = moderately low; 4 = moderate/average; 5 = moderately high; 6 = high; 7 = very high. Note that you will be using your own concept or image of what the ideal form is for each category. Before you rate the examples in each category, read the brief description, and read over all the examples within that category, getting a sense of how well the various objects conform to the category ideal. If you don't know an object that is listed, just put a question mark beside that object name, and move on to the next item.
>
> For each category, there may be additional items from your own list which you produced in the previous task. If you listed items for that category in the previous task, but they are not in the provided list, add them to the provided list in

the available spaces for that category. Once you have done this, rate each item's degree of conformity to the category ideal. Also rate any new items you may have added. Once you've completed one category, go to the next one and apply the same procedure, and so on. Complete one category at a time. You may ask any questions at any time. Please let the experimenter know when you are finished.

Shape Typicality Rating Information and Agreement

The mean Pearson correlations between judges' shape typicality ratings (Js-Js) were significant (all two-tailed $ps < .001$) but were lower than desirable, as shown in Table A1. This indicates variation between judges, possibly in how individual object items are construed and in differing conceptions of what constitutes an ideal for the category. However, the mean correlations between judges and the average typicality ratings (Js-Avg) appear to be adequately stable.

Table A1

Typicality Means and Mean Agreement

Comparison		*df*	Mean	*r*
Overall	Js-Avg	345	5.18	.80
	Js-Js	338		.51
FR Main List	Js-Avg	61	5.42	.82
	Js-Js	60		.58

Notes. Js = judges; Avg = average rating of judges' ratings; *df* = mean degrees of freedom (rounded off to the nearest whole number) for Pearson correlations; Mean = mean typicality rating. Note that ratings from two of the six judges were excluded due to evidence of response set (e.g., a judge making a long run of circling 6s, despite variability in the example items) and inconsistencies. The ratings from the remaining four judges were used here. The *r* values are the mean Pearson correlations between four judges (Js-Js) or between the four judges and the average rating (Js-Avg).

Appendix B

Instructions for FE Participants' Exemplar Specification, Object Shape Grouping, and Similarity Rating

Exemplar Shape Specification (First Domain)

During the transcription, the experimenter put a question mark beside any name that referred to an object that was ambiguous within the task context. Specific exemplar object information was obtained from the participant's drawing and more specific name or description. Exemplar specification instructions were as follows:

> Here (indicating the participant's transcription sheet) is your list of __________ objects. For names with a question mark, if you can remember the object you thought of, draw its basic shape outline. For example, if you said *candy*, draw the basic shape–just a quick sketch–of the one that you thought of. Draw the shape right beside the object name (experimenter demonstrates). Write the more specific name of the kind of object (e.g., _____________ or _____________ for candy). Write the specific name just underneath the object name, in the same slot (experimenter demonstrates). If you can't remember the shape or the specific kind of object, write another question mark, and then draw the shape and write the specific name of the object that is most familiar to you. You can ask me questions at any time. Please let me know when you are finished. You may start now.

Object Shape Grouping, First Domain: Instructions

The object shape grouping instructions were as follows:

> The task is to categorize the _________ objects according to their basic three-dimensional shape. The first thing to do is read over your list to get an idea of the different kinds of three-dimensional shapes of the objects. To fully appreciate the three-dimensional shape of an object, it is necessary to visualize it from multiple angles, as if you are picking it up and looking at it. You will notice that the three-dimensional forms of the objects may differ between each other in thickness, proportion of length to width, curvature versus straightness, shape flexibility, and various structural complexities. Ignore other aspects like the orientation or size of the object, its material composition, its conceptual category (e.g., whether it is a food, sports item, etc.), and other factors. Focus only on shape.
>
> When you find two or more items together that have the same basic kind of shape, bracket them like this (experimenter demonstrates on an example sheet).

Then draw the basic three-dimensional shape, and write a letter (e.g., K) for that group. For each new category you find, draw the new shape and write the new letter. Label single unique items by the same method. When you find items with shapes that you've already identified, bracket and then label them with the category letter used previously. Note that each shape you draw is not intended to be a precise depiction of an object's shape; it is just an "icon" having the approximate shape of the category that you have in mind. When you find items that have a kind of shape that you've already identified, bracket and then label them with the category letter used previously. The shape classifications should be based on the specific objects as you thought of them during the listing task.

When you're grouping the objects, you may find that some items are easier to classify than others. For the difficult ones that don't seem to fit clearly into one group, skip past them temporarily. You may later find that they can be grouped with other items. On the other hand, some items may be unique in that they don't belong with any category. Don't try to force them into a category where they don't belong. In those cases, label the unique object with its own letter and shape drawing.

When you've finished grouping your objects, check your groupings. You can make changes, such as merging or splitting categories, but only if you think they'll improve the classification.

Once you've checked everything over and have made any needed changes, you may find that some of your categories are similar to each other. Draw lines linking together those categories that are similar to each other. The categories that you link should be clearly similar. Use no more than two links per category.

For all the categories that have only one or two items, draw lines linking each to its most similar category or categories.

Here are the guidelines (see below) in case you need to refer to them. Start by reading over your list to survey the kinds of three-dimensional shapes, then do the grouping. You can ask questions at any time. Please let me know as soon as you're done.

Guidelines for Object Shape Grouping

The guidelines sheet summarizing the instructions was also available for the participant throughout the task:

1. Read over the list, getting an idea of the kinds of 3-dimensional object shapes.

2. Bracket and label the shape categories for single items and groups of items.

3. You can make changes at any time, such as relocating items, or merging or splitting some categories, if you think these changes will improve your classification.

4. Draw lines linking different categories that are similar to each other. Use no more than two (2) links per category. Categories that are not clearly similar to each other should not be linked.

5. Link categories that have only one (1) or two (2) items to their most similar other category.

6. Tell the experimenter as soon as you've finished.

Shape Similarity Rating, First Domain.

In this task, using the output transcript, and having considered the objects' shapes in the previous task, the participant rated the shape similarity between the adjacent pairings of list items. The experimenter demonstrated the rating procedure using the example sheet, while delivering the instructions, as follows:

> This will be a shape similarity rating task. In the previous task, you assigned objects to shape categories, but this time you will give a more precise judgement of the overall shape similarity between pairs of objects. For each sequential pair of objects in the list, you'll rate the similarity between the three-dimensional forms of the objects. Use "7" for pairs with a very high level of shape similarity, and "1" for pairs with a very low level of shape similarity. (Experimenter shows the rating scale. The rating scheme for degree of similarity is as follows: 1 = very low; 2 = low; 3 = moderately low; 4 = moderate; 5 = moderately high; 6 = high; 7 = very high. Experimenter shows an example of how the pairs of objects are rated). Making accurate ratings generally requires making use of all the numbers on the scale, from 1 to 7. The ratings within your list of _________ object pairs should be logically consistent with each other (show example, using participant's output). Focus only on the shape similarity between the objects. (Review shape features: thickness, proportion of length to width, curvature versus straightness, shape flexibility, and other complexities). Ignore other features like size, orientation, conceptual category, or material. The rating should be based on the objects that you thought of during the task. If you can not remember the specific object, make your similarity rating based on the shape of the object most familiar to you. First read over the list and make some shape similarity ratings mentally. Then, once you have a general idea of how you would rate some of the object pairs, start putting down your ratings in pencil. When you've rated all the pairs, check them over. It's important to correct or adjust any ratings that you think need to be changed in order to improve their accuracy. You can ask questions at any time. Let me know when you're finished.

For the second domain, the instructions for the exemplar specification, object shape grouping, and shape similarity rating were the same as those above except the wording was much abbreviated. Again, the guidelines and rating scales, and the experimenter, were available for the participant throughout the procedure.

Appendix C

Shape Classification of Master Sets

Instructions for Shape Classification Card Sort of a Master Set

The instructions for classifying a domain's master set of objects by shape were as follows:

> This set of cards has the names of a large number of objects. All of the objects in this set had been listed by other participants who completed another task where they listed ________ objects. The cards are in random order. [The experimenter had shuffled and mixed the cards thoroughly into random order and then spread them around on the table haphazardly in front of the participant]. The task will be to sort the cards into separate columns according to the different specific kinds of three-dimensional shapes of the objects. Objects that have the same or similar kind of three-dimensional shape should be grouped together in the same column. Make a different column for each different type of three-dimensional shape, like this (experimenter demonstrates). Take into account the various aspects of three-dimensional shape, including curvature versus straightness, proportion of length to width, thickness, shape flexibility, and other structural shape complexities that an object might have. In classifying the objects, ignore the size, material, or other aspects. Focus only on object shape.
>
> As you sort the objects into groups, you may find that some items are easier to classify than others. For the difficult ones that don't seem to fit clearly into one group, put them aside temporarily. You may later find that they will fit into a group. Otherwise, if they don't fit, leave them to the side. If you don't know the shape of a particular object, let me know. You may recognize the object if it is described, even if you don't recognize the name used here.
>
> Once you have sorted the cards, check your groupings over. You can make changes to your groupings, joining some and splitting others, if you think this will improve the classification.
>
> Once you've finished classifying the objects by shape, let me know.
>
> Next, you'll sort your *groupings* according to shape. Stack the cards for each grouping to keep them together within their group. For categories that are most similar to each other in some shape feature, have the stacks touching each other directly, like this (experimenter demonstrates). The categories that are not similar should not touch directly.
>
> You may refer to these instructions (experimenter hands participant the sheet, see below). You can ask questions at any time. Please let me know as soon as you are finished. This task may take close to the full two hours.

Guidelines for Shape Category Sorts

The summary instructions for object shape sorting were available to the participant for reference during the task, and were written on a sheet as follows:

1. Sort object items into separate columns according to basic 3-dimensional shape.

2. Put uniquely-shaped objects to one side, and come back to them later.

3. If you don't know the shape of an object, notify the experimenter.

4. Check over your groupings. Make changes if they will improve the classification.

5. Stack the items in each category.

6. Sort the stacks according to shape similarity. Stacks that are similar should be touching each other directly; those that are not similar should not touch directly.

Appendix D

Semantic Classification of Master Sets

Instructions for Semantic Classification Card Sort of a Master Set

The procedure, verbal instructions, and written guidelines for this task were similar to those for the shape master list sort, except for the reference to conceptual (semantic) categories. The uniquely conceptual (semantic) categorical subject matter in the instructions is excerpted here:

> …The task will be to sort the cards into different columns according to the different general conceptual category of the object items. For example, some things may be kinds of fruit or food, money, kitchen utensils, furniture, or any of a large number of different conceptual categories. Ignore physical aspects of the objects, such as shape, size, or material. Classify the items only according to the conceptual categories…

Appendix E

Size Classification of Master Sets

Object Size Card Sort of a Master Set

Size was rated in a sorting task that used the same items as used in the shape and semantic master set sorts. Extremely small (e.g., axon) and extremely large (planet) objects were excluded from the participant's sorting procedure. Such items were assigned by default to the extreme smallest and largest categories, respectively. The instructions for the object size sort were as follows:

> The deck of cards that you have contains the names of __________ objects. The cards are in random order. The task will be to sort these _________ objects into 9 groups according to the typical size of the objects, from smallest to the largest. Before you form any groups, put about 60 cards out onto the table and read them over, getting an idea of the variety and range of object sizes in the list. To form an estimate of overall size for each object, take into consideration how long the object is from one end to the other, and the surface area and volume of the object. Next, start to sort these cards into 9 groups, from the group with the smallest objects on the left, to the group with the largest objects on the right. Then take more cards and add to your groups. You can make adjustments by moving objects from one size group to another, if you think this will improve the size classification. If you have difficulty deciding the size group in which to place a certain object card, put that card off to one side temporarily and come back to it later. If you don't know a particular object, let me know. You may recognize the object from a description, even if you don't recognize the name used here. Once you have formed your groups, check them over to make sure no object items are in the wrong size group. If a very large number of objects is concentrated in one group, you may wish to divide the group into two parts to distinguish between smaller and larger objects in that group. You may refer to a written set of guidelines that summarize the instructions for this task (hands participant Guidelines sheet; see next page). You can ask me questions at any time. Also, please let me know when you are finished. The first task may take about half an hour to 45 minutes.

Instructions were reviewed, very briefly, for the size sort of the second domain.

Guidelines for Object Size Sort

The task guidelines were written on a sheet and were available for the participant throughout the task:

1. Sort the objects into 9 size groups, from smallest (on the left) to largest (right).

2. In estimating overall size, take into consideration the object's length, surface area, and volume.

3. Read over several cards to get an idea of the relative sizes of the objects, and begin to assign objects to the nine size categories. Then add more cards to the appropriate size groups.

4. If you don't know the size of an object, notify the experimenter.

5. You can move objects from one size group to another at any time, if this will improve the size classifications.

6. When you have assigned all of the objects to size categories, check them over. If one group contains a very large number of objects, it may be possible to divide the group into two parts to separate the smaller versus larger objects in that category.

Appendix F

Formulae for Measures of Categorical Clustering

Table F1

Categorical Clustering Measures: ARC and MRR

Terms	Formulae
n = number of items reported in output	
	Max = n – c
n_i = n of items output in the *i*th category	
	ARC: $E(r) = (\Sigma n_i^2 / n) - 1$
c = number of categories in output	
	$ARC_{obtained} = [r - E(r)] / [Max - E(r)]$
r = number of categorical repetitions[a]	
	$ARC_{chance} = 0$
Max = maximum possible r for an output	
	$MRR_{obtained} = r / Max$
E(r) = chance expected r for an output	
	$MRR_{chance} = [(E - 1)(n - 1)] / [(Max)(N - 1)]$
E = number of items per category in the study list	
N = number of items in the study list	

Notes. ARC scores range from a minimum of –1 to a maximum of +1, with a chance value of 0 (zero). MRR scores range from a minimum of 0 (zero) and a maximum of 1.

Appendix G

Adequacy of Recall Levels for Valid Measurement of Clustering

Murphy and Puff (1982) wrote that "...an accurate assessment of clustering depends on having a reasonably large number of items recalled from several categories..." (p. 121). Yet they did not expand on how to ascertain that reasonably large number. I will attempt to do so here. *First*, a categorical repetition cannot possibly occur unless two (2) or more items from a category are in fact recalled. Thus, the minimum number of items recalled should be at least two per category for most of the categories in the list. Obtaining an average of two items recalled per category on a single study-recall trial, in FR, is often not possible (but see below). *Second*, the total number of items recalled should require people to use mostly higher organizational processes, such as categorical grouping. A person can recall most of the items by rote if the total number of items to be recalled doesn't exceed short-term memory (STM) capacity by a considerable margin. *Third*, clustering scores at low levels of recall tend to exhibit high variability. Practically speaking, the lower the levels of recall, the larger the samples of participants needed to obtain valid, reliable means and distributions of clustering scores.

Concerning the total items recalled, if we assume that the average adult's STM capacity for words is between five (Graesser & Mandler, 1978) and seven (Miller, 1956) units, then recall should be considerably more than this if the researcher is interested in higher organizational schemes such as categorical clustering. Clustering obtained when the number of items recalled is mostly within STM capacity may not reflect the same underlying organizational processes that occur when the number of items recalled exceeds STM capacity. If the researcher wants to ensure that as many or more of the items were recalled according to higher organizational schemes than by rote, this puts the minimum level of recall at between about 10 and 14 items (i.e., an average of about two items per mnemonic unit, for five to seven such units). The minimum could be even higher, depending on how many categories of items are in the to-be-recalled list.

These above considerations are especially important where the researcher wishes to test the null hypothesis of chance clustering. Consider a to-be-recalled list of 20 items presented with 5 exemplars in each of 4 categories, where item order is randomized. According to the two-items-per-category criterion, minimum recall is eight (8) items. According to the STM capacity five to seven mnemonic-units criterion, the minimum is between 10 and 14 items. Unfortunately, it is often difficult to obtain these mean levels of recall on the first trial, where the study list had been presented just once. Generally, this means that to test the null hypothesis with sufficient confidence, the researcher must use multiple study-test trials, or use multiple presentations of the list before a single recall trial, or use some other manipulation to increase recall. Otherwise, if there is insufficient recall, a null result could simply mean that not enough items were recalled for clustering to occur.

Appendix H

Zimmer's (1989) Shape Clustering and Recall Results

Table H1

Zimmer's (1989) Free Recall Experiment 1: Nouns

	Instructions			
	Standard		Imagery	
Measure	Trial 1	Trial 2	Trial 1	Trial 2
Proportion Recalled	.36	.58	.29	.48
ARC	-.04	.13	-.06	-.02

Summary. Recall and clustering results are presented for two free recall trials. Participants were tested individually. Participants in one group were instructed to recall the words (standard instructions; n = 10) and those in the other group were instructed to generate mental imagery of each object (imagery instructions; n = 10) to be recalled. The list of 32 object names contained four exemplars in each of eight shape categories, plus four buffer items (two at each end of the list) that were not included in scoring. Shape categories included spherical, triangular, rod/post-like, flat rectangular, block-like, ring-like, disks, and string-like. Words were presented aurally in randomized unblocked sequences (i.e., no same-shape pairings) and were recalled in standard written format. Note that measures of variability were not reported in Zimmer, (1989).

(See next page for Zimmer's shape clustering results for Experiments 2 and 3).

Table H2

Zimmer's (1989) Free Recall Experiment 2: Pictures

	Instructions	
Measure	Standard	Imagery
Proportion Recalled	.74	.68
ARC	.12	.16

Summary. Displayed are recall and clustering scores for free recall of object pictures, with standard and imagery instruction groups (n = 10 per group). Slide photos of the objects in their typical orientations were presented in randomized unblocked sequences and were recalled by name in writing. There were 4 exemplars in each of six shape categories, plus six buffer items (three at each end of the list) and six filler items not included in scoring. Shape categories included board-like, block-like, spheres, cylinders, rings, and disks. (Buffers and fillers were of objects of different shape categories). The list was presented twice before a single recall trial.

Table H3

Zimmer's (1989) Free Recall Experiment 3: Nouns

	Instructions	
Measure	Standard	Imagery
Proportion Recalled	.51	.51
ARC	-.01	.05

Summary. Recall and ARC scores are shown for free recall of object nouns, with standard and imagery instructions groups (overall n = 26). The procedure followed that of Experiment 1, but there was no second study-recall trial. In addition, Experiment 3 included a cued recall test (not shown here) after the free recall. The list of 28 object names contained seven shape categories with four exemplars per category, plus four buffer items (two at the each end of the list) not included in scoring.

Appendix I

Free Recall Lists

Table I1

List 1 Object Nouns and Shape Categories

Spheroid	Long Pointed Cyls.	Flat Rectangular	String-like
pearl	pencil	placemat	rope
marble	toothpick	chessboard	snake
volleyball	arrow[a]	tile	licorice
globe	nail	card	wire
bubble	thorn	diskette	hose

Note. Cyls. = Cylinders.
[a]For two participants, the word *spear* was used instead of *arrow*.

Table I2

List 2 Object Nouns and Shape Categories

Long Cylinders	Hollow Cones	Block-Like	Disks
steel pipe[a]	funnel	dumpster	plate
chopstick	pylon	casket	hubcap[a]
pool cue	teepee	juice box	frisbee
cigarette[a]	witch's hat	brick	CD
log	snowcone	train car	quarter

Note. [a]For the first two participants exposed to List 2, *flagpole*, *wand*, and *button* (not shown here) were used. Because one of those two participants organized those three items with others by first letter, those three items were replaced, for all subsequent List 2 participants, by *steel pipe*, *cigarette*, and *hubcap*.

Table I3

List 3 Object Nouns and Shape Categories

Medium Cylinders	String/Strip	Flat Rectangular	Rings
marker	leash	envelope	Life Saver/Cheerio
cigar	ribbon/yarn	shingle	diamond r./nosering
test tube	whip/dental floss	licence plate	bike tire
cannon	eel/worm	playing card	wreath
flute	cable	mat	hula hoop

Notes. The forward slash separating alternative words indicates that, in those cases, one word (e.g., *diamond ring*) was used for some of the participants, and the other word was used for the other participants (e.g., *nosering*). The more frequently used alternative is listed on the left side of the forward slash. Two participants received the first version of List 3, which had Slab Rectangular (**book, tombstone, cutting board, mattress, domino*) instead of Flat Rectangular, and Bulbs (**light bulb, pear,* **doorknob, bowling pin, teardrop*) instead of Medium Cylinders. In addition, for Rings, those two participants had *bracelet* instead of *diamond ring*, *steering wheel* instead of *bike tire*, and for String/Strip-like had **shoelace* instead of *cable*. *These items were construed as objects in the room (or, in the case of *book*, as an object that could be in the room), and grouped on that basis, by the first two participants to receive the first version List 3. Therefore, the researcher replaced those items with others for main version of List 3 for subsequent participants. Note that for light bulb, the actual light in the room was the long straight fluorescent tube type. Later, another participant in the List 3 group made mnemonic analogies between a *diamond ring* and a different ring that she was wearing, and between *marker* and a pen in the room, and between *mat* and the wall-to-wall carpeting in the room. However, she did not group these items together in the card sort but arranged them as separate individual units each of Size 1.

Appendix J

FR Item Characteristics

Normative information for many of the items with respect to several item characteristic variables was not available. Relevant word association strength information was obtained for a subset of the items. This latter information is reported in the FR Results section. Typicality, word length, and object size information is reported here. Note that in cases where alternate words were used (e.g., *worm* or *eel* in List 3, see Appendix I, Table I3), the mean typicality, mean word length, and mean rated object size of the objects (e.g., of *worm* and *eel*) were used in the analysis of the items in the main FR list.

The mean shape typicality levels (where 1 = very low and 7 = very high) for items included in the main FR lists ($M = 5.42$, $SD = 1.076$) were slightly higher than those from the rated set that were not included ($M = 5.13$, $SD = 1.088$), $t(354) = -1.945$, $SE = 0.1488$, $p = .053$ (two-tailed). The inter-judge correlations for shape typicality ratings are shown in Appendix A. Mean shape typicality did not differ among the 12 main list FR categories, $F(11, 48) = 1.33$, $MSE = 0.795$, $p = .236$, or the three main FR lists, $F(2, 57) = 1.10$, $MSE = 0.842$, $p = .34$.

Word length, as measured by the number of letters in the word, did not differ between categories [$F(11, 48) = 1.46$, $MSE = 4.225$, $p = .18$] or lists [$F(2, 57) = 0.61$, $MSE = 4.651$, $p = .549$]. (There remained no significant differences in word length for categories or lists when the object name *CD*—scored as having only two letters—was excluded from the analyses). When word length was measured by the number of syllables in the word, there were again no significant differences for categories or lists.

The object size ratings provided by judges in the previous FE classification study were not strictly applicable to the items used in this study because of the different rating contexts. (For an example of a discrepancy that arose when those ratings were applied here to the FR items, *witch's hat*, originally rated by size in the round master set, had a larger size rating than *casket*, which was rated in the straight master set). Nevertheless, it was important to examine the differences, within lists, between shape categories in mean object size. Therefore, I made size ratings of the objects using a rating scale of 1 to 9. There were 55 items in the main FR list that each had an average size rating from the ratings by the FE study size judges. The average size ratings by the FE size judges for those 55 main FR list objects were correlated r (53) = .82 with my size ratings. Using my size ratings on the main list items ($n = 60$), the analysis showed that there were some significant differences in object size between the 12 categories, $F(11, 48) = 2.32$, MSE = 2.396, $p = .02$. Tukey HSD post hoc tests showed that there were two significant differences: The block rectangles were larger than the long pointed cylinders ($p = .006$) and the spheres ($p = .02$), respectively. No other differences between categories were significant. Note that the block rectangles were used in List 2, but the spheres and the long pointed cylinders were used in List 1. That is, there were no significant size differences between categories within lists. There were significant differences between lists, $F(2, 57) = 4.61$, $MSE = 2.661$, $p = .01$. Specifically, List 2 objects had larger size ratings than List 1 objects ($p = .01$) according to the Tukey HSD test. No other differences between lists were significant.

Appendix K

Free Recall (FR) Instructions

First Trial

The instructions before the first trial list presentation were as follows:

> In the first trial of the task, you will hear the names of numerous real physical things or objects. Then, when all of the object names have been presented, you'll count backwards for two minutes. After that, I'll turn on the tape-recorder and ask you to say the names of as many of the objects as you can recall. Once you've done the recall, that will be the end of the first of four of these trials. You'll do each of the trials in the same sequence: You'll hear the words, then count backwards, then recall the object names. To aid your concentration, I'd like you to close your eyes, starting now, and keep them closed until all of the words have been read. The order of the words does not matter. The task is simply to recall as many of the object names as possible, in any way you wish. As you listen, when you hear each object name, think of that object. Ready? I'll pause for a few seconds and then I'll start.

After the list was presented, the experimenter gave the instructions for the distracter interval task:

> Okay, open your eyes. Now the task will be to count backwards by threes (3s) from 400 (or other number on later trials), out loud, for the next two minutes. The goal is to concentrate on counting accurately and continuously. I'll let you know when two minutes is up. Ready? Start.

When the participant completed the two minutes of backwards counting, the experimenter gave the recall instructions: "Alright, keeping your eyes closed again, say the names of as many of the objects as you can recall." A recall trial was ended when, after the participant had reported several items and seemed to have exhausted what he or she could recall, and 15 s elapsed since the last item's report, the prompt "Can you recall any more?" elicited a "no," or else prior to the prompt the participant indicated that she/he could recall no more. The audio recorders were shut off at the end of the trial.

(See next page for subsequent trial instructions).

Subsequent Trials

Encoding: "Okay, now I'll present the objects' names again. Close your eyes. Ready?"

Distracter Task: "Open your eyes. Count backwards by threes from 600 (or other number) out loud."

Recall: "Close your eyes, and say the names of as many objects as you can recall."

After the end of last recall trial: "Okay, open your eyes. That's the end of the recall experiment. You can take off the microphone. For the next part, I have some questions about your thinking and memory processes that occurred during the experiment."

Appendix L

Randomizations for MRR Chance Estimates for Low Levels of Recall

I found that there was a problem with the MRR chance formula (see Appendix F). Specifically, the formula does not produce accurate chance estimates when there are 5 to 7 items recalled from the 20 item list (four categories with five exemplars per category) used in the present study. For example, for a recall sequence of five items that contains four categories, the MRR chance formula produced a score of .84, which was unrealistically high. Indeed, .84 suggests near-perfect (1.0) clustering. To obtain valid estimates of chance MRR, I took actual random samples of items from the 20 item (4 categories x 5 exemplars) list. For each of sample sizes of 5, 6, or 7 items, I took an average of approximately 90 random samples from the randomized 20 item (4 categories x 5 exemplars) list. In most cases the random samples contained four categories, but some had three categories. I then found the mean MRR score obtained (r / n – c) for the random samples of the different sizes (5, 6, and 7 items) and numbers of categories (3 or 4). I then used those means as the chance MRR scores for the appropriate recall outputs.

To check the concurrent validity of these random sampling estimates of the MRR chance scores for the first trial outputs that had less than 8 items recalled (n = 12), I computed the correlation between the MRR difference scores (MRR obtained minus MRR random sampling estimate of chance) and the ARC scores for those outputs. The correlation was r(10) = .94. For that same subsample (n = 12), using the original chance MRR formula (MRR obtained minus MRR original chance calculation), the correlation between MRR difference scores and ARC scores was r(10) = .82, indicating a loss of accuracy for the original MRR chance formula for low levels of recall. Thus, the random sampling estimate of MRR chance for outputs having less than eight items recalled produces scores that are consistent with those produced by the well-known, well-established ARC measure. Therefore, I used the random sampling estimate of chance MRR for the outputs having less than eight items recalled.

I also examined random samples from the 20 item (4 categories x 5 exemplars) list that had 8, 10, 12, or 16 items. In those cases, the original MRR chance formula scores were consistent with MRR scores obtained from the random samples. For the subsample of first trial outputs having eight or more items recalled (n = 21), the correlation between the MRR difference scores (using the original formula to calculate MRR chance) and ARC was $r(19) = .985$, indicating a good level of accuracy. Therefore, I used the MRR chance formula for recall sequences of 8 or more items. Nevertheless, researchers should be advised to use random sampling for estimating valid chance values for MRR when recall is low.

Appendix M

Formulae for the Measure of Subjective Organization

Table M1

Subjective Organization Measure: SO-ARC

Terms	Formulae
M = number of items recalled on trial *t*	Max = M – 1 – R
N = number of items recalled on trial *t* + 1	E(ITR) = 2 Max / N
ITR = intertrial repetition; a pairing of items on trial *t* that are also paired on trial *t* + 1	$\text{SO-ARC}_{obtained}$ = [O(ITR) – E(ITR)] / [Max – (ITR)]
R = number of pairs from trial *t* that contain items *not* recalled on trial *t* + 1	SO-ARC_{chance} = 0
Max = maximum possible number of pairwise bidirectional ITRs	
[a]O(ITR) = obtained number of pairwise bidirectional ITRs	
E(ITR) = chance expected number of pairwise bidirectional ITRs	

Notes. [a]For example, if trial t's sequence is FGEBCAD, and trial t + 1's sequence is GACBDFE, there are two pairwise bidirectional ITRs. For further details, see Murphy (1979) and Murphy and Puff (1982).

Appendix N

Object Noun Master Sets

Notes on the Object Classifications

Caveat. The lists below do not constitute an exhaustive sampling of everyday objects. The items were obtained from a small number of university students who completed a limited set of free emission tasks, and were classified by a limited number of previously-untrained judges within relatively brief classification sessions. The lists below were not intended to be used for normative purposes, but are intended for use within the present studies. Nevertheless, researchers who are interested in these classifications are welcome to contact the researcher for further information.

The Coding of the Object Sets. Above the listing of each object set (below) there are general shape and general semantic classification labels. Whereas the classification labels provided within the lists (Tables N1-4) show the specific level of classification, the general or broader level classification labels are provided prior to each list. The general level classifications were based on participants' groupings of similar specific categories (i.e., sorting of stacks, as described in the instructions for master list sorting, Appendix C). The general classifications capture the judges' mergers of specific categories. Items classified as unique (i.e., not included in categories of two or more objects) are labelled under shape category by object name, or by an abbreviation or some distinctive label. Objects for which a clear size judgement could not be obtained, or for which a size judgement was problematic (e.g., extremely large objects), are labelled with "-----." Note that some objects of extreme size were assigned to higher than 9 or less than 1.

Classification Decisions in the Scoring of FE List Output Clustering. As was mentioned in the Methods section regarding the scoring of inter-item transitions as either cluster, switch, or filler, it was necessary to take into account information (such as shape similarity ratings, and the FE participant's classifications) in addition to that provided by the master list sorts below. This was important particularly because the classifications produced from the relatively small set of objects provided in an FE output list might not be the same as those produced from the larger master list sets in some cases. It must also be kept in mind that some specific exemplar information, provided by the FE participant's sketch-drawings of certain objects, may indicate slightly different forms than those possibly imaged by the readers based on the labels given below.

Long Objects

Shape Classification

FLSLRect (Flat Slab Rectangular): RecFlat1, RecFlat2, RectSlab.

LbdRect (Long Board-Like Rectangular): BoardSh, BoardLg, BrdLgThk.

LCyl (Long Cylinders): CylLong, CylLgThn, CylLgHlw, CylLngPt, CylLgRnd.

RcktPdll (Racket/Paddle): PaddleWd, PaddlReg, PaddlLac, PaddleSpo, PaddlSpa, PaddlDbl.

String: StrngThn, StrngThk.

Strip: StripThn, StripBrd.

Other Categories (not merged): BarRect, CylMed, Bottle, Tongue, PryCrow, Knife, Lizard, Boat, LongGun, RectSki, ArmLeg, DecidTree, Flat2DRc.

Semantic Classification

FoodBev: Food, FoodUtns, Beverage.

BldgMtPt (Building Materials, Parts): BldgPart, BldgMatr.

StatiArt (Stationery, Art-Related): Stationr, ArtRelat.

Other Categories (not merged): FurnAppl, Bldgs, Wood, ToolFstn, SewUtMat, Clothing, BathPers, BodyPart, Animals, PlantsRe, Sport, Weapon, Musical, Electric, PublUtil, Walking, Vehicle, String.

Table N1

Long Objects Master Set

Long Object	Shape	Semantic	Size
bench	BarRect	FurnAppl	7
candy bar (Mars)	BarRect	Food	2.5
carton	BarRect	carton	4.5
chimney	BarRect	BldgPart	8
drink box	BarRect	Beverage	2.5
eraser	BarRect	Stationr	1.5
fridge	BarRect	FurnAppl	7
power bar (electrical)	BarRect	Electric	3
semi's trailer	BarRect	Vehicle	9
skyscraper	BarRect	Bldgs	9
train	BarRect	Vehicle	9
carpet in hallway	Flat2DRc	BldgPart	8
driveway	Flat2DRc	PublUtil	8.5
field	Flat2DRc	Sport	9
highway	Flat2DRc	PublUtil	9
road	Flat2DRc	PublUtil	9
roadway or path	Flat2DRc	PublUtil	9
sidewalk	Flat2DRc	PublUtil	8.5
fence slat/picket	BoardLg	BldgPart	5.75
hardwood floor (board/strip)	BoardLg	BldgPart	5.5
handle on fridge	BoardLg	FurnAppl	4
shelf (one "board" of)	BoardLg	FurnAppl	5.5
siding	BoardLg	BldgPart	7

board	BoardLg	BldgMatr	6
see-saw	BoardLg	Sport	7
ruler	BoardLg	Stationr	3
yardstick/metre stick	BoardLg	Stationr	4.5
rail (flat rectangular)	BoardLg	BldgPart	6
joiner between flooring surfaces	BoardLg	BldgPart	4.5
gum (stick type)	BoardSh	Food	1.25
street name sign	BoardSh	PublUtil	4.5
running board (on truck)	BoardSh	Vehicle	5.75
chair leg	BrdLgThk	FurnAppl	4
2x4 of wood	BrdLgThk	BldgMatr	6
post (in a fence)	BrdLgThk	fencepst	6
shingle	RecFlat1	BldgPart	5.75
large window	RecFlat1	BldgPart	7
mirror	RecFlat1	BldgPart	5.5
glass (sheet)	RecFlat1	BldgMatr	6.5
piece of plywood	RecFlat1	BldgMatr	7.5
plexiglass	RecFlat1	BldgMatr	7
sheet of paper	RecFlat1	Stationr	3.5
poster (flat, on display)	RecFlat1	ArtRelat	5
framed art	RecFlat1	ArtRelat	5.75
picture	RecFlat1	ArtRelat	5.75
circuit board	RecFlat1	Electric	4
piece of cardboard	RecFlat1	pcardbrd	5.5
bill (money)	RecFlat1	moneybll	2.5
boxes flattened	RecFlat1	boxsfltg	5.5
solar panel	RecFlat2	solpanel	6
cupboard door	RecFlat2	FurnAppl	5.5
table	RecFlat2	FurnAppl	7
table top	RecFlat2	FurnAppl	7
door	RecFlat2	BldgPart	7
electrical outlet cover	RecFlat2	Electric	2.5
light switch cover	RecFlat2	Electric	2.5
box of Smarties	RectSlab	Food	2.5
candy bar	RectSlab	Food	2.5
calculator	RectSlab	Stationr	2.5
harmonica	RectSlab	Musical	2.5
remote control	RectSlab	Electric	2.5
slice of bread	RectSlab	Food	3
ski	RectSki	Sport	5.5
surfboard	RectSki	Sport	6.5
canoe	Boat	Sport	7.5
kayak	Boat	Sport	7.5
boat	Boat	Sport	9
rifle	LongGun	Weapon	5
bazooka	LongGun	Weapon	6.5

baseball bat	CylBottl	Sport	5
beer bottle	CylBottl	Beverage	3.5
glass bottle	CylBottl	Beverage	3.5
straw	CylLgHlw	FoodUtns	2
pipe (copper)	CylLgHlw	BldgMatr	4.5
pipe (e.g., household)	CylLgHlw	BldgMatr	5.5
PVC pipe	CylLgHlw	BldgMatr	6
bamboo	CylLgHlw	Wood	4.5
flute	CylLgHlw	Musical	4
conduit	CylLgHlw	Electric	5
culvert	CylLgHlw	PublUtil	8
tube (rigid)	CylLgHlw	tube	3.5
clarinet	CylLgHlw	Musical	4.5
recorder	CylLgHlw	Musical	4
wiener	CylLgRnd	Food	2.5
cucumber	CylLgRnd	Food	3.5
french bread	CylLgRnd	Food	4
body of airplane (fuselage)	CylLgRnd	Vehicle	9
fishing rod	CylLgThn	Sport	5.5
car antenna	CylLgThn	Vehicle	4
CB antenna	CylLgThn	Vehicle	4
carrot	CylLngPt	Food	2.5
toothpick	CylLngPt	FoodUtns	1
signaling tower	CylLngPt	Bldgs	9
screw	CylLngPt	ToolsFst	1
nail	CylLngPt	ToolsFst	1.25
spike	CylLngPt	ToolsFst	2
pick (straight bar)	CylLngPt	ToolsFst	5.5
drilling rod	CylLngPt	ToolsFst	5
needle (sewing)	CylLngPt	Sewing	1
pin	CylLngPt	Sewing	1
knitting needle	CylLngPt	Sewing	2.75
golf tee	CylLngPt	Sport	1
javelin	CylLngPt	Sport	5.5
arrow	CylLngPt	Weapon	4
jousting lance	CylLngPt	Weapon	6.5
missile	CylLngPt	Weapon	7
pen	CylLngPt	Stationr	2
pencil	CylLngPt	Stationr	2
pointer	CylLngPt	Stationr	4
drumstick (for drum)	CylLngPt	Musical	3
grounding rod	CylLngPt	Electric	5
oscilloscope probe	CylLngPt	oscllscp	3
artist's paint brush	CylLngPt	ArtRelat	2.75
wand (conductor's)	CylLngPt	Musical	3
magic wand	CylLngPt	magiwand	3

CN tower	CylLngPt	Bldgs	9
cinnamon stick	CylLong	Food	2
light pole	CylLong	PublUtil	8
steel pole	CylLong	steelpl	6.5
chopstick	CylLong	FoodUtns	2.25
office bulb (tube)	CylLong	BldgPart	5
closet bar	CylLong	BldgPart	5
railing1	CylLong	BldgPart	6
rail on staircase	CylLong	BldgPart	6.5
stick	CylLong	Wood	4
handle (e.g., of broom)	CylLong	ToolsFst	5.5
broomstick	CylLong	ToolsFst	5.5
tree trunk	CylLong	PlantsRe	7.5
basketball net post	CylLong	Sport	6.5
volleyball post	CylLong	Sport	6.5
piece of chalk	CylLong	Stationr	2
pole	CylLong	PublUtil	6.5
telephone pole	CylLong	PublUtil	8
cigarette	CylLong	cigarett	1.5
flagpole	CylLong	flagpole	7
tree branch	CylLong	Wood	6
walking stick	CylLong	Walking	5
street lamp	CylLong	PublUtil	8
screwdriver	CylMed	ToolsFst	3
paper towel roll	CylMed	FoodUtns	4
piece of firewood	CylMed	Wood	5
shampoo bottle	CylMed	BathPers	3.5
marker	CylMed	Stationr	2.25
test tube	CylMed	testtube	2.25
glass (tall, drinking)	CylMed	Beverage	3.5
jug	CylMed	Beverage	4.5
juice container	CylMed	Beverage	4.5
Q-tip	PaddlDbl	BathPers	1.5
kayak paddle	PaddlDbl	Sport	6
shovel	PaddleSpo	ToolsFst	5.5
spoon	PaddleSpo	FoodUtns	2.5
badminton racket	PaddleWd	Sport	5
tennis racket	PaddleWd	Sport	5
violin	PaddleWd	Musical	5
guitar	PaddleWd	Musical	5.75
guitar case	PaddleWd	Musical	5.75
lacross stick	PaddlLac	Sport	5.5
oar	PaddlReg	Sport	6
paddle	PaddlReg	Sport	5.5
spatula (spreader)	PaddlSpa	FoodUtns	3
spatula (flipper)	PaddlSpa	FoodUtns	3

fruit roll-up	StripBrd	Food	2.5
boarder material	StripBrd	BldgPart	5
scarf	StripBrd	Clothing	4.5
sod (laid out)	StripBrd	PlantsRe	5.75
paper banner	StripBrd	paperbnr	5
measuring tape (tape portion)	StripThn	ToolsFst	4
ribbon	StripThn	Sewing	3.5
belt	StripThn	Clothing	4.5
whip	StripThn	Weapon	5
dog leash	StripThn	dogleash	4
tube (unspecified, flexible)	StrngThk	tube	3.5
hose	StrngThk	ToolsFst	5
dog's tail	StrngThk	BodyPart	3.5
person's spine	StrngThk	BodyPart	5
elephant trunk	StrngThk	BodyPart	7
slug	StrngThk	Animals	1.25
centipede	StrngThk	Animals	1.5
worm	StrngThk	Animals	1.5
snake	StrngThk	Animals	4
eel	StrngThk	Animals	4.5
beanstalk (mythic, "Jack and…")	StrngThk	PlantsRe	9
rope	StrngThk	String	5
licorice	StrngThn	Food	2.25
spaghetti noodle	StrngThn	Food	2
thread	StrngThn	Sewing	1
coat hanger	StrngThn	Clothing	4
shoelace	StrngThn	Clothing	3.5
hair (strand)	StrngThn	BodyPart	1
fishing line	StrngThn	Sport	2
jump/skipping rope	StrngThn	Sport	5
cable	StrngThn	Electric	4
coaxial cable	StrngThn	Electric	4.5
cord	StrngThn	Electric	4
extension cord	StrngThn	Electric	4
house wiring	StrngThn	Electric	4.5
wire (unspecified)	StrngThn	Electric	3
telephone/power line	StrngThn	PublUtil	7
string	StrngThn	String	2
telephone cord	StrngThn	Electric	4
sock	Tongue	Clothing	3.5
tongue	Tongue	BodyPart	2.5
arm	ArmLeg	BodyPart	5
elephant leg	ArmLeg	BodyPart	7
leg	ArmLeg	BodyPart	5.5
finger	ArmLeg	BodyPart	2.25
tree1 (palm)	DecidTree	PlantsRe	9

tree2 (deciduous)	DecidTree	PlantsRe	9
prybar	PryCrow	ToolsFst	4
crowbar	PryCrow	ToolsFst	4.5
breadknife	Knife	FoodUtns	2.5
butterknife	Knife	FoodUtns	2.5
knife (small handle)	Knife	FoodUtns	2.5
knife (unspecified)	Knife	FoodUtns	2.5
sword	Knife	Weapon	5
scissor blade	Knife	Stationr	3
lizard (gecko)	Lizard	Animals	3
alligator	Lizard	Animals	7
airplane	airplane	Vehicle	9
baseboard (quarter round)	basebrdq	BldgMatr	6
bathtub	bathtub	BathPers	7
bedrail on pick-up truck	bedrailp	Vehicle	5.5
bobby pin	bobbypin	BathPers	1
bridge	bridge	PublUtil	9
cane	cane	Walking	5
celery stick	celrystk	Food	2.25
couch	couch	FurnAppl	7
crutch	crutch	Walking	5.5
curtain rod	curtrod	BldgPart	5
drapes	drapes	BldgPart	7
faucet	faucet	BldgPart	3
fingernail	fingnal	BodyPart	1
flower (rose, one stem)	flower	PlantsRe	3
flower vase	flowrvs	PlantsRe	4.5
fork	fork	FoodUtns	2.5
giraffe neck	giraffen	Animals	7.5
golf club	golfclub	Sport	5
hairbrush	hairbrus	BathPers	2.5
hammer	hammer	ToolsFst	3.5
handsaw	handsaw	ToolsFst	4.5
hockey stick	hckystck	Sport	5.5
ladder	ladder	ToolsFst	7.5
ladle	ladle	FoodUtns	3
limosine	limosine	Vehicle	8.25
Oh-Henry bar	ohhenryb	Food	2.5
paper clip	paperclp	Stationr	1
pocket comb	pocktcb	BathPers	2.5
railway	railway	PublUtil	9
rake	rake	ToolsFst	5.5
roadguard	roadgurd	PublUtil	8
saxophone	saxphon	Musical	5
slide	slide	Sport	7.5
toothbrush	toothbru	BathPers	2.25

tower (electrical)	tower	PublUtil	9
tree3 (spruce)	treecon	PlantsRe	9
trumpet	trumpet	Musical	4.5
utility knife	utilknif	ToolsFst	2.5
windshield wiper	wndwipr	Vehicle	4
wrench	wrench	ToolsFst	3

Flat Objects

Shape Classification

Disks: Disks, DisksThk, DisksCrv.

FLRSLR (Flat and Slab Rectangles): RectThn1, RctThnFlx, RectThn2, RctThnFlx2.

SlabThk (Thick Slab Rectangles): RctSlbFlx, RectSlab.

Other Categories (not merged): Paddle, Shoe, Knife, Strips, Fish, Kite, Triangl, Rlong, Table, Ski, Flat2D.

Semantic Classification

FoodReKU (Food-Related, Kitchen Utensils): Food, FoodUtns.

BldgPtMt: BldgPart, BldgMatr.

Other Categories (not merged): SportGam, Weapons, ArtPictu, ElectrEn, Furnitur, ClothLin, Bathroom, Plants, Animals, BodyPart, Optics, ToolsFst, PocketTh, Vehicle, Water, Street, Musical, Stationr, Electric.

Table N2

Flat Objects Master Set

Flat Object	Shape	Semantic	Size
burger flattener	Disks	FoodUtns	3.00
button (on shirt)	Disks	ClothLin	1.00
CD	Disks	ElectrEn	3.00
coin	Disks	PocketTh	1.75
cookie	Disks	Food	2.25
cymbal	Disks	Musical	5.00
dime	Disks	PocketTh	1.00
DVD	Disks	ElectrEn	3.00
frisbee	Disks	SportGam	5.00
garbage can lid	Disks	FoodUtns	5.50
lillypad	Disks	Plants	3.75
loonie	Disks	PocketTh	1.75
manhole cover	Disks	Street	6.50
mat1	Disks	BldgPart	6.50
nailhead	Disks	ToolsFst	0.50
nickel	Disks	PocketTh	1.75
paint palette	Disks	ArtPictu	4.50
pancake	Disks	Food	4.25
penny	Disks	PocketTh	1.00
pizza	Disks	Food	5.50
plate	Disks	FoodUtns	5.00
quarter	Disks	PocketTh	1.75
tortilla shell	Disks	Food	4.00
twonie	Disks	PocketTh	1.75
capacitor	DisksCrv	ElectrEn	0.50
landmine	DisksCrv	Weapons	4.50
lens of glasses	DisksCrv	Optics	2.25
lens of magnifying glass	DisksCrv	Optics	2.25
rock (rounded, flat)	DisksCrv	rockFlat	2.50
shield	DisksCrv	Weapons	6.50
bottle cap	DisksThk	FoodUtns	1.75
box of chocolates	DisksThk	Food	5.00
Cert	DisksThk	Food	1.00
clock (circular)	DisksThk	clock	4.00
poker/casino chip	DisksThk	SportGam	1.75
puck	DisksThk	SportGam	3.25
screwhead	DisksThk	ToolsFst	0.50
tuning knob on radio	DisksThk	ElectrEn	1.00
flat fish (unspecified)	FishSh	Animals	5.00

flat fish (with stinger)	FishSh	Animals	5.00
front lawn	Fl2DRect	Plants	9.50
patio	Fl2DRect	BldgPart	9.00
wall	Fl2DRect	BldgPart	9.00
kite	KiteSh	SportGam	6.00
stingray	KiteSh	Animals	6.50
butterknife	KnifeSh	FoodUtns	3.00
knife1 (dinner)	KnifeSh	FoodUtns	3.00
machete	KnifeSh	Weapons	5.50
saw	KnifeSh	ToolsFst	5.50
sword	KnifeSh	Weapons	6.00
badminton racket	Paddle	SportGam	5.50
frying pan	Paddle	FoodUtns	5.00
key	Paddle	PocketTh	1.50
ping pong paddle	Paddle	SportGam	4.25
snowshoe	Paddle	SportGam	5.50
tennis racket	Paddle	SportGam	5.50
bed (mattress and Boxspr)	RctSlbFlx	Furnitur	8.50
mattress	RctSlbFlx	Furnitur	8.50
paddle float	RctSlbFlx	SportGam	6.00
phone book	RctSlbFlx	Stationr	5.25
school agenda	RctSlbFlx	Stationr	4.25
wallet	RctSlbFlx	PocketTh	3.50
aluminum foil	RctThnFlx	FoodUtns	4.50
bill (money)	RctThnFlx	PocketTh	2.50
blanket	RctThnFlx	ClothLin	8.50
carpet	RctThnFlx	BldgPart	9.00
construction paper	RctThnFlx	Stationr	5.50
curtain	RctThnFlx	BldgPart	8.00
envelope	RctThnFlx	Stationr	3.75
flag	RctThnFlx	flag	6.50
flooring	RctThnFlx	BldgPart	9.00
garbage bag	RctThnFlx	FoodUtns	6.50
linoleum	RctThnFlx	BldgPart	-----
paper (standard sheet)	RctThnFlx	Stationr	4.50
plastic sheeting	RctThnFlx	BldgMatr	8.00
postage stamp	RctThnFlx	Stationr	1.00
rug	RctThnFlx	BldgPart	9.00
Saran wrap	RctThnFlx	FoodUtns	4.50
sheet (bed sheet)	RctThnFlx	ClothLin	8.00
table cloth	RctThnFlx	ClothLin	8.00
tin foil	RctThnFlx	FoodUtns	4.50
towel	RctThnFlx	ClothLin	6.50
wax paper	RctThnFlx	FoodUtns	4.50
welcome mat	RctThnFlx	BldgPart	6.50
Ziploc bag	RctThnFlx	FoodUtns	3.00

breadslice	RctThnFlx2	Food	3.50
brochure	RctThnFlx2	Stationr	4.25
magazine	RctThnFlx2	Stationr	5.00
mousepad	RctThnFlx2	ElectrEn	4.50
newspaper	RctThnFlx2	Stationr	5.00
oven mit	RctThnFlx2	FoodUtns	4.50
brick	RectBlock	BldgMatr	3.50
cabinet top	RectLong	Furnitur	7.50
counter top	RectLong	Furnitur	8.25
fence slat	RectLong	BldgMatr	6.50
gum (stick type)	RectLong	Food	1.75
lumber	RectLong	BldgMatr	7.00
mantle	RectLong	BldgPart	7.00
piece of board (long)	RectLong	BldgMatr	7.00
piece of wood (2x4)	RectLong	BldgMatr	6.00
ruler	RectLong	Stationr	3.00
see-saw	RectLong	SportGam	8.00
shelf	RectLong	Furnitur	7.00
sidewalk (one section)	RectLong	Street	8.50
slide	RectLong	SportGam	7.50
step	RectLong	BldgPart	5.50
strip of hardwood floor	RectLong	BldgPart	6.50
tailgate on truck	RectLong	Vehicle	7.50
top of bar (for serving drinks)	RectLong	Furnitur	8.25
1'x1' cement slab	RectSlab	BldgMatr	5.50
audio cassette	RectSlab	ElectrEn	3.50
bar of soap (hotel)	RectSlab	Bathroom	2.25
bar of soap2	RectSlab	Bathroom	2.75
binder	RectSlab	Stationr	5.25
binder (full)	RectSlab	Stationr	5.25
book1	RectSlab	Stationr	5.00
box of Smarties	RectSlab	Food	3.00
calculator	RectSlab	Stationr	3.00
candy bar	RectSlab	Food	3.00
cell phone	RectSlab	ElectrEn	3.50
cementslab2	RectSlab	BldgMatr	6.00
chair seat	RectSlab	Furnitur	5.50
cutting board	RectSlab	FoodUtns	5.50
domino	RectSlab	SportGam	1.75
DVD player	RectSlab	ElectrEn	5.50
Etch-a-sketch	RectSlab	SportGam	5.50
flat screen TV	RectSlab	ElectrEn	6.50
flat screen display/monitor	RectSlab	ElectrEn	5.50
giftbox	RectSlab	giftbox	5.50
Jenga block	RectSlab	SportGam	2.00
keyboard	RectSlab	ElectrEn	5.00

laptop computer	RectSlab	ElectrEn	5.75
LCD display (monitor)	RectSlab	ElectrEn	5.75
luminescence display	RectSlab	ElectrEn	1.75
project box	RectSlab	Electric	5.00
scoreboard	RectSlab	SportGam	7.50
Sony digital recorder	RectSlab	ElectrEn	3.00
textbook	RectSlab	Stationr	5.50
VCR	RectSlab	ElectrEn	5.50
VHS/video tape cassette	RectSlab	ElectrEn	4.00
bank card	RectThn1	PocketTh	2.25
board game	RectThn1	SportGam	5.50
bus pass	RectThn1	PocketTh	2.25
circuit board	RectThn1	ElectrEn	5.00
cooking pan	RectThn1	FoodUtns	5.00
credit car	RectThn1	PocketTh	2.25
glass pane	RectThn1	BldgPart	7.50
glasssheet	RectThn1	BldgMatr	7.50
ID card	RectThn1	PocketTh	2.25
index card	RectThn1	Stationr	3.00
licence plate	RectThn1	Vehicle	4.50
magnet	RectThn1	magnet	2.00
mirror	RectThn1	BldgPart	6.00
pan	RectThn1	FoodUtns	5.00
picture (photo)	RectThn1	ArtPictu	3.50
placemat	RectThn1	FoodUtns	4.50
playing card	RectThn1	SportGam	2.25
plexiglass sheet	RectThn1	BldgMatr	7.50
project box cover	RectThn1	Electric	4.50
putty palette	RectThn1	ToolsFst	4.50
razor blade (paint scraper)	RectThn1	ToolsFst	1.75
sheet of metal	RectThn1	sheetMet	-----
steel plate	RectThn1	metalPla	4.50
tray	RectThn1	FoodUtns	5.00
trowel for plaster	RectThn1	ToolsFst	4.25
window	RectThn1	BldgPart	7.50
CD case	RectThn2	ElectrEn	3.75
clipboard	RectThn2	Stationr	5.00
door	RectThn2	BldgPart	8.25
DVD case	RectThn2	ElectrEn	3.75
end (tab) of zipper	RectThn2	ClothLin	0.50
end of slot screwdriver	RectThn2	ToolsFst	0.50
floppy disk 3.5"	RectThn2	ElectrEn	3.00
gyprock (drywall, sheet)	RectThn2	BldgMatr	8.50
key chain tab	RectThn2	PocketTh	1.75
kitchen tile	RectThn2	BldgMatr	3.75
notebook/exercise book	RectThn2	Stationr	5.00

painting	RectThn2	ArtPictu	5.50
picture frame	RectThn2	ArtPictu	5.50
plywood (sheet)	RectThn2	BldgMatr	8.50
ramp (toward doorway)	RectThn2	BldgPart	8.50
roof top	RectThn2	BldgPart	9.50
sheet of cardboard	RectThn2	sheetcrdb	6.00
shingle	RectThn2	BldgMatr	6.00
sign	RectThn2	Street	6.50
table top1	RectThn2	Furnitur	8.25
table top2	RectThn2	Furnitur	8.25
tile	RectThn2	BldgMatr	3.75
bottom/sole of shoe	ShoeSh	ClothLin	4.00
flip flop	ShoeSh	ClothLin	4.00
ski	SkiSledN	SportGam	6.50
snowboard	SkiSledN	SportGam	6.50
ironing board	SkiSledW	ClothLin	7.50
skateboard	SkiSledW	SportGam	5.50
surfboard	SkiSledW	SportGam	7.50
toboggan	SkiSledW	SportGam	7.50
rope (error)	String	ToolsFst	5.50
spaghetti (error)	String	Food	2.25
bacon	Strips	Food	2.50
belt	Strips	ClothLin	4.00
seatbelt	Strips	Vehicle	5.00
sock	Strips	ClothLin	3.75
strip/ribbon on lampshade	Strips	Furnitur	3.00
tapeworm	Strips	Animals	2.75
air hockey table	TableRct	SportGam	7.50
bench	TableRct	Furnitur	8.00
coffee table	TableRct	Furnitur	7.50
desk	TableRct	Furnitur	8.00
Japanese table	TableRct	Furnitur	8.25
piano stool	TableRct	Furnitur	7.00
picnic table	TableRct	Furnitur	8.50
ping pong table	TableRct	SportGam	8.50
pool table	TableRct	SportGam	8.50
table	TableRct	Furnitur	8.25
guitar pick	Triangul	Musical	1.75
slice of pizza	Triangul	Food	3.50
badge	badge	badge	2.50
carpenter's square	carpntrsq	ToolsFst	5.50
chassis of trailer	chassisT	Vehicle	-----
fingernail	fingrnal	BodyPart	1.00
floor (2D)	floor2D	BldgPart	9.00
fork	fork	FoodUtns	3.00

French bread (loaf; error)	frnchbrd	Food	5.00
goalie stick	goalstck	SportGam	6.50
graduation cap	gradcap	ClothLin	4.50
hood of a car	hoodcar	Vehicle	8.00
ice surface on pond (2D)	icesurf2D	Water	9.50
iron	iron	ClothLin	4.00
ladder	ladder	ToolsFst	8.50
LCD TV screen	lcdTVscr	ElectrEn	6.50
leaf	leaf	Plants	2.25
lipstick (error)	lipstck	Bathroom	1.75
bathmat	Oval	Bathroom	6.50
roof of a car	roofcar	Vehicle	8.50
screen (on a door or window)	screen	BldgPart	6.50
shirt	shirt	ClothLin	6.00
side view of pencil (2D)	sidviewp	Stationr	1.75
spatula (flipper)	spatulaF	FoodUtns	3.00
top of spatula (spreader)	spatuTop	FoodUtns	2.25
steak	steak	Food	3.75
stop sign	stopsign	Street	5.50
tongue	tongue	BodyPart	3.00
trunk of a car	trunkcar	Vehicle	8.00
washboard	washbord	ClothLin	6.25
water surface (2D)	watersrf2D	Water	9.50

Round Objects

Shape Classification

Spherical/Spheroid: Spheroid, SphCmplx, SphLong, SphLong2, Sphdntpl, SphTaper, SphIrreg, SphSemi.

Bulbs/Bottles: Bulb, BulbBott.

Cylinders: CylindM, CylindL1, CylindKn, CylindLr, Cylindsh, CylCupSh, CylCup, CylindBt

Long Cylinders: CylindL2, CylindL3, CylindLp.

Bowls: Bowl, BowlElon.

Disks: Disk, DiskVFl, DiskTh2, DiskTh3, DiskRnd, DiskTapr.

Rings: Rings, RingFlat, RingBand, RngThFlx, RingThck, RingFram, RingElon, RingElFr

Other Categories (not merged): SpiralFl, String, Fltrdrop, CircLn2D, Circ2DRg, Circ2DDk.

Semantic Classification

FoodBv: Fruit, Vegtbls, Bredmeat, Sweets, Foodcont, Bevergs, Condimnt.

AutoVehS: Vehicle (airplane), Automotv, Bike, Street.

Other Categories (not merged): Kitchenu, Medphysi, Opticvis, Lights, Sewknitt, Jewellry, Pocketth, Weapons, SafetyRe, Sprtgame, Bathpers, Bodypart, Pets, ShapeNam, SymblPic, Vegplant, Stationr, Applelec, Bldgpart, Entelect, Musical, Clothing, Geologic, Containr, Toolmech, Celestia, Yard.

Table N3

Round Objects Master Set

Round Object	Shape	Semantic	Size
ball bearing	Spheroid	Toolmech	0.75
ball in computer mouse	Spheroid	Entelect	1.75
ball joint (for trailer)	Spheroid	Automotv	3.50
ball of chalk	Spheroid	Sprtgame	4.00
ball of wool	Spheroid	Sewknitt	4.25
ball of yarn	Spheroid	Sewknitt	4.25
ball on earring stud	Spheroid	Jewellry	1.00
ball point of ball point pen	Spheroid	Stationr	0.50
baseball	Spheroid	Sprtgame	3.50
basketball	Spheroid	Sprtgame	5.75
beach ball	Spheroid	Sprtgame	6.00
bead	Spheroid	Jewellry	1.00
bee-bee	Spheroid	Weapons	0.50
blueberry	Spheroid	Fruit	1.00

bouncing ball (small, rubber)	Spheroid	Sprtgame	2.50
bubble	Spheroid	Water	2.00
cannon ball	Spheroid	Weapons	5.00
chicken ball	Spheroid	Bredmeat	2.75
dogberry	Spheroid	Fruit	1.00
doorknob2	Spheroid	Bldgpart	3.50
Earth (extra large)	Spheroid	Celestia	-----
eyeball	Spheroid	Bodypart	2.50
faucet/tap handle	Spheroid	Bathpers	2.75
glass dome (at Disney)	Spheroid	Sprtgame	9.50
globe	Spheroid	Celestia	6.00
gobstopper	Spheroid	Sweets	1.75
golf ball	Spheroid	Sprtgame	2.50
grapefruit	Spheroid	Fruit	3.75
gumball	Spheroid	Sweets	1.75
hamster ball	Spheroid	Pets	4.50
little bubble in bubble tea	Spheroid	Bevergs	0.50
magic 8-ball	Spheroid	Sprtgame	3.75
marble	Spheroid	Sprtgame	1.50
medicine ball	Spheroid	Sprtgame	6.00
micelle (extra small)	Spheroid	Medphysi	-----
microphone top	Spheroid	Entelect	2.75
moon (extra large)	Spheroid	Celestia	-----
musketball	Spheroid	Weapons	2.00
orange	Spheroid	Fruit	3.50
pea	Spheroid	Vegtbls	1.00
pearl	Spheroid	Jewellry	1.50
ping pong ball	Spheroid	Sprtgame	2.50
planet (extra large)	Spheroid	Celestia	-----
pool ball	Spheroid	Sprtgame	3.50
rock1(round)	Spheroid	Geologic	2.75
scoop of ice cream	Spheroid	Sweets	3.50
snowball	Spheroid	Sprtgame	3.50
soccerball	Spheroid	Sprtgame	5.75
softball	Spheroid	Sprtgame	3.50
sphere	Spheroid	ShapeNa	-----
star (extra large)	Spheroid	Celestia	-----
street hockey ball	Spheroid	Sprtgame	3.50
stressball	Spheroid	Sprtgame	3.00
sun (extra large)	Spheroid	Celestia	-----
tangerine	Spheroid	Fruit	3.50
tennis ball	Spheroid	Sprtgame	3.50
the world (extra large)	Spheroid	Celestia	-----
volleyball	Spheroid	Sprtgame	5.50
head (person's)	SphCmplx	Bodypart	5.25
cantaloupe	SphLong	Fruit	5.00

grape	SphLong	Fruit	1.75
honeydew melon	SphLong	Fruit	5.00
kiwi	SphLong2	Fruit	3.50
lemon	SphLong2	Fruit	3.50
lime	SphLong2	Fruit	3.50
watermelon	SphLong2	Fruit	5.75
candy	SphLong2	Sweets	1.75
football	SphLong2	Sprtgame	5.00
apple	Sphdntpl	Fruit	3.50
plum	Sphdntpl	Fruit	3.25
tomato	Sphdntpl	Fruit	3.50
onion	SphTaper	Vegtbls	3.50
turnip	SphTaper	Vegtbls	5.00
egg	SphTaper	Bredmeat	3.25
Easter egg	SphTaper	Sweets	2.75
Kinder surprise	SphTaper	Sweets	3.25
balloon	SphTaper	Sprtgame	5.25
potato	SphIrreg	Vegtbls	3.50
motorcycle helmet	SphSemi	SafetyRe	5.75
fish bowl	SphSemi	Pets	5.00
light fixture	SphSemi	Bldgpart	5.00
pear	Bulb	Fruit	3.50
light bulb	Bulb	Lights	3.50
bell	Bulb	Musical	4.75
drop of water	Bulb	Water	0.50
beer bottle	BulbBott	Bevergs	4.00
bottle of pop	BulbBott	Bevergs	4.25
Pepsi bottle	BulbBott	Bevergs	4.25
wine bottle	BulbBott	Bevergs	4.50
bowling pin	BulbBott	Sprtgame	5.00
can of soup	CylindM	Foodcont	4.25
canned food	CylindM	Foodcont	4.25
can of Pepsi	CylindM	Bevergs	4.25
can of pop	CylindM	Bevergs	4.25
drink can	CylindM	Bevergs	4.25
thermos	CylindM	Kitchenu	4.50
cylinder	CylindM	ShapeNa	-----
battery	CylindM	Entelect	2.50
capstan on guitar (for string)	CylindM	Musical	1.00
barrel	CylindM	Containr	7.00
keg	CylindM	Containr	7.00
oil drum	CylindM	Containr	7.50
Pringle chips container	CylindLI	Sweets	4.50
Rollo chocolate (whole roll)	CylindLI	Sweets	3.25
roll of paper towel	CylindLI	Kitchenu	5.00
telescope	CylindLI	Opticvis	6.25

roller (for hair)	CylindL1	Bathpers	2.50
log	CylindL1	Vegplant	8.00
silo	CylindL1	Containr	9.00
water tower	CylindL1	Containr	9.00
mushroom	CylindKn	Vegtbls	2.50
flashlight	CylindKn	Lights	4.00
doorknob (standard)	CylindKn	Bldgpart	3.50
microphone	CylindKn	Entelect	2.75
gear shift	CylindKn	Automotv	4.00
handle (e.g., on a bike)	CylindKn	bike	3.50
screwdriver	CylindKn	Toolmech	3.25
bread loaf	CylindLr	Bredmeat	5.00
Subway sub	CylindLr	Bredmeat	4.50
finger	CylindLr	Bodypart	2.50
stereo	CylindLr	Entelect	6.00
airplane	CylindLr	Vehicle	9.00
marshmallow	Cylindsh	Sweets	2.50
roll of toilet paper	Cylindsh	Bathpers	4.25
tree branch	CylindL2	Vegplant	5.75
closet rod	CylindL2	Bldgpart	5.75
pillar	CylindL2	Bldgpart	8.00
exhaust pipe	CylindL2	Automotv	5.75
lever to control wipers, etc.	CylindL2	Automotv	2.75
street pole	CylindL2	Street	8.00
fence pole	CylindL2	Yard	7.00
drill bit	CylindL2	Toolmech	1.75
cigarette	CylindL2	cigarett	2.50
metal pipe	CylindL2	metalpip	5.00
dancing pole	CylindL2	dancingp	7.00
straw	CylindL3	Bevergs	2.50
cannula (extra small)	CylindL3	Medphysi	-----
spoke	CylindL3	bike	2.50
syringe	CylindLp	Medphysi	2.50
needle	CylindLp	Sewknitt	1.00
javelin	CylindLp	Sprtgame	7.00
pen	CylindLp	Stationr	2.50
nail	CylindLp	Toolmech	1.75
screw	CylindLp	Toolmech	1.75
cup2 (child's)	CylCupSh	Kitchenu	3.50
pot	CylCupSh	Kitchenu	5.75
yoghurt container	CylCup	Foodcont	4.25
cup1	CylCup	Kitchenu	4.25
cup3	CylCup	Kitchenu	4.25
glass	CylCup	Kitchenu	4.25
jar	CylCup	Kitchenu	4.25
beaker	CylCup	Medphysi	4.00

medication bottle	CylCup	Medphysi	2.75
thimble	CylCup	Sewknitt	1.75
dryer drum	CylCup	Applelec	7.00
washer drum	CylCup	Applelec	7.00
pail	CylCup	Containr	5.25
salt shaker	CylindBt	Condimnt	2.75
coffee press	CylindBt	Kitchenu	4.50
bullet	CylindBt	Weapons	1.75
toothpaste tube	CylTaper	Bathpers	3.25
bowl	Bowl	Kitchenu	4.50
dish	Bowl	Kitchenu	4.50
drain (catch piece)	Bowl	Kitchenu	3.25
contact lens	Bowl	Opticvis	1.00
shin pad	Bowl	SafetyRe	4.25
satellite dish	Bowl	Entelect	8.00
roast pan	BowlElon	Kitchenu	6.00
Skydome	BowlElon	Sprtgame	10.00
bathtub	BowlElon	Bathpers	7.50
limited slip differential	Disk	Automotv	4.50
pancake	Disk	Bredmeat	4.75
pita bread	Disk	Bredmeat	4.50
pizza	Disk	Bredmeat	6.00
coaster	Disk	Kitchenu	3.00
cookie sheet	Disk	Kitchenu	6.00
pan	Disk	Kitchenu	6.00
plate	Disk	Kitchenu	5.00
platter	Disk	Kitchenu	5.50
glasses lens	Disk	Opticvis	2.75
frisbee	Disk	Sprtgame	5.00
trampoline	Disk	Sprtgame	8.00
thumbtack top	Disk	Stationr	1.00
mirror	Disk	Bldgpart	6.25
window	Disk	Bldgpart	6.25
CD	Disk	Entelect	4.25
DVD	Disk	Entelect	4.25
record	Disk	Entelect	5.25
hubcap	Disk	Automotv	5.75
OZ FM car sticker	DiskVFl	ozfmcars	3.75
cookie	DiskTh2	Sweets	2.75
circular tea bag	DiskTh2	Bevergs	3.00
button (of shirt)	DiskTh2	Jewellry	1.00
coin	DiskTh2	Pocketth	1.75
dime	DiskTh2	Pocketth	1.00
loonie	DiskTh2	Pocketth	2.00
nickel	DiskTh2	Pocketth	1.75
penny	DiskTh2	Pocketth	1.50

quarter	DiskTh2	Pocketth	1.75
twonie	DiskTh2	Pocketth	2.00
checker	DiskTh2	Sprtgame	2.00
poker chip	DiskTh2	Sprtgame	2.00
weight plate	DiskTh2	Sprtgame	5.75
turntable	DiskTh2	Entelect	5.75
manhole cover	DiskTh2	Street	6.50
pulley (wheel)	DiskTh2	Toolmech	4.50
top of a stool	DiskTh2	tpfstool	6.00
bottle cap	DiskTh3	Bevergs	1.75
lid	DiskTh3	Kitchenu	3.00
tea light candle	DiskTh3	Lights	2.50
puck	DiskTh3	Sprtgame	3.50
sandbox	DiskTh3	Sprtgame	7.00
animal's water bowl	DiskTh3	Pets	4.50
thermostat dial	DiskTh3	Applelec	2.50
earphone piece	DiskTh3	Entelect	1.50
headphone earpiece	DiskTh3	Entelect	2.00
stereo dial	DiskTh3	Entelect	2.25
drum	DiskTh3	Musical	5.50
tire2	DiskTh3	Automotv	6.50
clock	DiskTh3	clock	5.00
Smartie	DiskRnd	Sweets	1.00
camera lens	DiskRnd	Opticvis	2.50
bumper boat	DiskRnd	Sprtgame	7.50
discus	DiskRnd	Sprtgame	5.00
soap1	DiskRnd	Bathpers	3.50
rock2 (flat)	DiskRnd	Geologic	2.75
pie	DiskTapr	Sweets	5.25
Reese's peanut butter cup	DiskTapr	Sweets	2.75
stovetop burner	SpiralFl	Applelec	5.00
thermostat coil	SpiralFl	Applelec	2.50
measuring tape	SpiralFl	Toolmech	3.50
shoelace hole	Rings	Clothing	0.75
wrist band (athletic)	Rings	Sprtgame	3.50
bangle	Rings	Jewellry	3.25
earring (hoop type)	Rings	Jewellry	2.00
nose ring	Rings	Jewellry	1.50
ring	Rings	Jewellry	1.50
key chain ring	Rings	Pocketth	2.00
basketball hoop rim	Rings	Sprtgame	6.00
hula hoop	Rings	Sprtgame	6.75
washer (fastener)	RingFlat	Toolmech	1.50
watch	RingBand	Jewellry	3.25
watch bracelet	RingBand	Jewellry	3.25
bracelet	RngThFlx	Jewellry	3.25

necklace	RngThFlx	Jewellry	3.25
bagel	RingThck	Bredmeat	3.75
Cheerio	RingThck	Bredmeat	1.00
doughnut	RingThck	Sweets	3.75
Life Saver	RingThck	Sweets	1.25
life-preserver	RingThck	SafetyRe	6.50
car tire	RingThck	Automotv	6.50
bearing (whole unit)	RingThck	Toolmech	2.50
nut (fastener)	RingThck	Toolmech	1.00
ferris wheel	RingFram	Sprtgame	9.50
merry-go-round1	RingFram	Sprtgame	8.00
merry-go-round2	RingFram	Sprtgame	8.00
hamster running wheel	RingFram	Pets	4.50
bike wheel	RingFram	bike	6.50
steering wheel	RingFram	Automotv	6.00
bike gear	RingFram	bike	4.25
belt buckle	RingElon	Jewellry	3.25
picture frame	RingElon	Bldgpart	6.00
tennis racket head	RingElFr	Sprtgame	5.75
axon (extra small)	String	Medphysi	-----
blood vessel (extra small)	String	Medphysi	-----
surgical tube	String	Medphysi	3.00
copper wire	String	Bldgpart	2.50
wiring	String	Bldgpart	2.50
cable cord	String	Entelect	3.00
telephone cord	String	Entelect	2.50
guitar string	String	Musical	2.50
hose	String	Yard	4.50
chain link fence	String	Yard	-----
flower petal	Leaf	Vegplant	1.75
leaf	Leaf	Vegplant	2.50
skating rink	CircLn2D	Sprtgame	9.50
track	CircLn2D	Sprtgame	9.50
roundabouts in Europe	Circ2DRg	Street	9.50
bottom of coffee cup	Circ2DRg	Kitchenu	3.50
belly button (outline of)	Circ2DRg	Bodypart	1.00
letter "o"	Circ2DRg	SymblPic	-----
zero	Circ2DRg	SymblPic	-----
traffic light (outline)	Circ2DRg	Street	4.25
bull's eye (outline)	Circ2DDk	Sprtgame	1.75
pupil	Circ2DDk	Bodypart	1.00
watch face	Circ2DDk	Jewellry	1.75
circle	Circ2DDk	ShapeNa	-----
smiley face	Circ2DDk	SymblPic	-----
knee brace	kneebr	SafetyRe	4.25
basketball netting	bsktnet	Sprtgame	5.75

fuse	fuse	Applelec	1.75
spring	spring	Toolmech	3.50
conveyer belt	convbelt	Toolmech	7.50
guitar soundhole	guitsndh	Musical	4.25
toilet bowl	toiltbwl	Bathpers	6.75
bulldozer scoop (error)	blldzscp	Toolmech	8.00
loofa	loofa	Bathpers	3.75
cave (round example)	cave	Geologic	8.00
hurricane	hurrican	hurrican	-----
banana	banana	Fruit	3.50
jack (e.g., for headphones)	jackaudi	Entelect	1.25
cupholder in car	zcuphold	Automotv	4.00
pylon (FR)	FR-Cones	Street	6.00
snowcone (FR)	FR-Cones	Sweets	4.25
witch's hat (FR)	FR-Cones	Clothing	5.75
wreath (FR)	FR-Cones	wreath	6.00

Note. Items marked "(FR)" were not generated by the FE participants, but were added for the purpose of obtaining classification information for the subsequent FR experiment.

Straight Objects

Shape Classification

BCLPk (Block-Cuboid, Block-Long, Block-with-Peak): BlckCube, BlcLong1, BlockPk.

SRthick (Slab thick Rectangular): SlabThck, SlabFlxb, SlabLong.

FRSRthin (Flat and thin Slab Rectangular): SlabThn1, SlabThn2, FlatThin, FlatFlex, Ribbed (flat rectangular).

LbdSRLbd (Long Board, Slab Rectangular Long Board): BlcLong3, BoardLng.

VLbdExtB (Very Long Board, Extremely long Block): BlcLong2, windshield wiper.

TableLike: TableStr, BenchLg.

Other Categories (not merged): 2DRect (Rect2D), Stair, Triangle, Knife, CarTruck, Frames, Cage, Cylindr (Cylinder, beveled).

Semantic Classification

FoodBvCU (Food, Food Containers and Utensils, and Beverages): Food, FoodCont, FoodDisp, FoodUtns.

Drug: DrugRecr, DrugMed.

Applianc: ApplncsK, ApplncsL.

Furnitur: FurnBdrm, FurnGen.

BldgPtMt: BldgMatr, BldgPart, BldgOut.

Building: Bldgs, BldgPubl, BldgPyrm.

ElectEnt (Electronics and Entertainment): ElctrCmpt, ElctrnAV, ElctrAcc, ElctrCom.

TranVehP: Vehcl, StrtSign.

Other Categories (not merged): Tools, PersHygn, ClothngR, SportGam, Stationr, Money, Death, Music, Symbolic, Picture, Garbage, Animal, Boxes.

Table N4

Straight Objects Master Set

Straight Object	Shape	Semantic	Size
bench	BenchLg	FurnGen	4.75
park bench	BenchLg	BldgPubl	5.00
side bench	BenchLg	SportGam	6.00
AC adapter	BlckCube	ElctrAcc	1.75
amplifier	BlckCube	ElctrnAV	4.00
box (general, unspecified)	BlckCube	Boxes	4.00
cardboard box1	BlckCube	Boxes	4.00
child's block	BlckCube	SportGam	1.75
die	BlckCube	SportGam	1.00
dishwasher	BlckCube	ApplncsK	5.50

dryer	BlckCube	ApplncsL	5.50
Game Cube	BlckCube	ElctrCmpt	3.50
garbage can2	BlckCube	Garbage	4.25
ice cube	BlckCube	Food	1.25
monitor (standard, not "flat")	BlckCube	ElctrCmpt	4.00
night table	BlckCube	FurnBdrm	4.75
safe	BlckCube	Money	4.75
stove/oven	BlckCube	ApplncsK	5.50
sugar cube	BlckCube	Food	1.00
TV	BlckCube	ElctrnAV	4.25
washing machine	BlckCube	ApplncsL	5.50
aquarium	BlcLong1	aquarium	4.50
boat ("cabin" of large vessel)	BlcLong1	Vehcl	8.00
bookshelf1	BlcLong1	FurnGen	5.00
bookshelf2	BlcLong1	FurnGen	5.50
box of tissue	BlcLong1	PersHygn	3.00
breadbox	BlcLong1	FoodCont	3.75
brick	BlcLong1	BldgMatr	2.75
building	BlcLong1	Bldgs	9.00
bus (metro)	BlcLong1	Vehcl	7.00
cabinet	BlcLong1	FurnGen	4.75
candy bar	BlcLong1	Food	2.00
case of beer (12-pack)	BlcLong1	DrugRecr	3.75
casket	BlcLong1	Death	5.50
CD player1	BlcLong1	ElctrnAV	3.25
cement block	BlcLong1	BldgMatr	3.75
computer (mini tower)	BlcLong1	ElctrCmpt	4.00
cupboards	BlcLong1	FurnGen	5.00
drawer	BlcLong1	FurnGen	3.75
dresser1	BlcLong1	FurnBdrm	5.50
dresser2	BlcLong1	FurnBdrm	5.50
electric razor	BlcLong1	PersHygn	2.25
elevator	BlcLong1	BldgPart	6.50
filing cabinet	BlcLong1	FurnGen	5.25
freezer	BlcLong1	ApplncsK	5.50
fridge	BlcLong1	ApplncsK	5.50
garbage bin (dumpster)	BlcLong1	Garbage	6.00
garbage can1	BlcLong1	Garbage	4.25
garbage can3	BlcLong1	Garbage	4.25
ink cartridge	BlcLong1	ElctrCmpt	1.75
juice box	BlcLong1	FoodCont	2.25
locker	BlcLong1	locker	4.50
mailbox2	BlcLong1	Stationr	3.50

mailbox3	BlcLong1	Stationr	3.50
microwave	BlcLong1	ApplncsK	4.25
milk carton	BlcLong1	FoodCont	3.00
school	BlcLong1	Bldgs	9.00
shipping container box	BlcLong1	Boxes	6.25
shoebox	BlcLong1	ClothngR	3.25
shower	BlcLong1	BldgPart	5.50
stereo	BlcLong1	ElctrnAV	3.75
stereo speaker	BlcLong1	ElctrnAV	4.00
toaster	BlcLong1	ApplncsK	3.50
toaster oven	BlcLong1	ApplncsK	3.75
toolbox	BlcLong1	Tools	3.75
train car	BlcLong1	Vehcl	7.00
Tupperware container1	BlcLong1	FoodCont	2.75
TV stand	BlcLong1	FurnGen	4.50
vending machine	BlcLong1	FoodDisp	5.50
piano key	BlcLong2	Music	1.75
skywalk	BlcLong2	BldgPubl	8.00
post	BlcLong3	BldgPubl	6.50
barn	BlockPk	Bldgs	9.00
church	BlockPk	Bldgs	9.00
garage	BlockPk	Bldgs	8.50
house	BlockPk	Bldgs	9.00
house frame	BlockPk	BldgPart	8.50
shed	BlockPk	Bldgs	8.00
2x4 of wood	BoardLng	BldgMatr	4.25
windshield wiper (rubber strip)	BoardLng	Vehcl	2.75
bookmark	BoardLng	Stationr	2.00
diving board	BoardLng	SportGam	4.50
fence slat	BoardLng	BldgOut	4.25
floorboard	BoardLng	BldgMatr	4.50
level	BoardLng	Tools	3.25
lumber	BoardLng	BldgMatr	5.00
razor (blade)	BoardLng	PersHygn	1.25
ruler	BoardLng	Stationr	2.25
shelf1 (one "board" of shelf)	BoardLng	BldgPart	4.25
siding (one piece)	BoardLng	BldgMatr	4.50
stair (one tread piece)	BoardLng	BldgPart	4.25
stick gum	BoardLng	Food	1.00
street name sign	BoardLng	StrtSign	3.75
strip of hardwood flooring	BoardLng	BldgMatr	3.50
warf	BoardLng	warf	8.00
window sill	BoardLng	BldgPart	4.00
wooden plank	BoardLng	BldgMatr	4.50
baby's crib	Cage	FurnGen	5.50

cage	Cage	cage	-----
car	CarTruck	Vehcl	6.75
truck	CarTruck	Vehcl	6.75
light pole	Cylindr	BldgPubl	6.75
pencil	Cylindr	Stationr	1.75
area rug	FlatFlex	BldgPart	6.50
crazy carpet	FlatFlex	SportGam	4.25
envelope	FlatFlex	Stationr	2.75
flag	FlatFlex	Symbolic	4.75
floor mat	FlatFlex	BldgPart	4.25
ham slice	FlatFlex	Food	2.50
mat	FlatFlex	BldgPart	4.25
money (bill)	FlatFlex	Money	2.00
pamphlet	FlatFlex	Stationr	2.75
piece of paper	FlatFlex	Stationr	3.00
poster	FlatFlex	Picture	4.25
stamp (postage)	FlatFlex	Stationr	1.00
ticket	FlatFlex	ElctrnAV	1.25
towel	FlatFlex	PersHygn	4.25
business card	FlatThin	Stationr	2.00
computer screen	FlatThin	ElctrCmpt	3.50
credit card	FlatThin	Money	2.00
folder	FlatThin	Stationr	3.25
knife blade3 (rectangular)	FlatThin	Tools	1.75
licence plate	FlatThin	Vehcl	3.00
mirror	FlatThin	BldgPart	4.00
motherboard	FlatThin	ElctrCmpt	3.25
name tag	FlatThin	Stationr	1.75
pan	FlatThin	FoodUtns	3.50
photo	FlatThin	Picture	2.00
playing card	FlatThin	SportGam	2.00
shingle	FlatThin	BldgMatr	3.75
sign1	FlatThin	StrtSign	4.00
sign2	FlatThin	StrtSign	4.75
speed limit sign	FlatThin	StrtSign	4.00
theatre screen	FlatThin	ElctrnAV	7.50
window	FlatThin	BldgPart	4.50
chain link fence section	Frame	BldgOut	5.50
picture frame	Frame	Picture	3.75
knife1 (breadknife)	Knife	FoodUtns	2.50
knife2 (butterknife)	Knife	FoodUtns	2.00
basketball court	Rect2D	SportGam	8.00
ceiling	Rect2D	BldgPart	8.00
driveway	Rect2D	BldgOut	8.00
outline behind basketball net	Rect2D	SportGam	4.00
patio	Rect2D	BldgOut	7.50

tennis court	Rect2D	SportGam	8.00
vent cover	Ribbed	BldgPart	3.50
washboard	Ribbed	ApplncsL	3.75
bed	SlabFlxb	FurnBdrm	5.50
mattress	SlabFlxb	FurnBdrm	5.50
bar (for serving drinks)	SlabLong	FurnGen	5.50
bar (granola)	SlabLong	Food	2.00
box of spaghetti	SlabLong	FoodCont	3.00
calculator	SlabLong	Stationr	2.00
CD ROM unit	SlabLong	ElctrCmpt	3.00
cell phone	SlabLong	ElctrCom	2.00
counter top	SlabLong	FurnGen	4.75
digital voice recorder	SlabLong	ElctrCom	2.00
keyboard	SlabLong	ElctrCmpt	3.50
RAM chip	SlabLong	ElctrCmpt	1.00
remote control	SlabLong	ElctrnAV	2.50
sidewalk	SlabLong	BldgPubl	5.00
audio tape case	SlabThck	ElctrnAV	2.00
block of cheese	SlabThck	Food	2.25
book2	SlabThck	Stationr	3.25
box of cereal	SlabThck	FoodCont	3.25
box of cookies	SlabThck	FoodCont	3.25
briefcase	SlabThck	Stationr	3.75
cardboard box2	SlabThck	Boxes	4.00
china cabinet	SlabThck	FurnGen	6.00
eraser	SlabThck	Stationr	1.00
fusebox	SlabThck	ElctrAcc	3.75
game console	SlabThck	ElctrCmpt	3.50
Hall's cough drop	SlabThck	Food	1.00
mailbox1	SlabThck	Stationr	3.50
match box	SlabThck	DrugRecr	1.25
MP3 player	SlabThck	ElctrnAV	2.00
Playstation2	SlabThck	ElctrCmpt	3.50
sandbox	SlabThck	SportGam	5.50
sandwich	SlabThck	Food	2.75
Sponge Bob Square Pants	SlabThck	ElctrnAV	-----
tape recorder	SlabThck	ElctrnAV	3.00
tombstone	SlabThck	Death	4.50
Tupperware container2	SlabThck	FoodCont	2.75
VCR	SlabThck	ElctrnAV	3.75
video tape (VHS) cassette	SlabThck	ElctrnAV	2.75
wallet	SlabThck	Money	2.00
X-Box	SlabThck	ElctrCmpt	3.50
binder	SlabThn1	Stationr	3.25
book1	SlabThn1	Stationr	2.75
CD player2	SlabThn1	ElctrnAV	2.75

cement slab	SlabThn1	BldgMatr	3.75
cigarette package	SlabThn1	DrugRecr	2.00
cutting board	SlabThn1	FoodUtns	3.25
domino	SlabThn1	SportGam	1.25
DVD player	SlabThn1	ElctrnAV	3.75
laptop computer	SlabThn1	ElctrCmpt	3.50
monitor ("flat")	SlabThn1	ElctrCmpt	3.75
pizza box	SlabThn1	FoodCont	3.75
textbook	SlabThn1	Stationr	3.25
base of microscope	SlabThn2	bsmicrsc	2.50
base of pool table (slate)	SlabThn2	SportGam	6.00
CD case	SlabThn2	ElctrnAV	2.75
chalkboard	SlabThn2	Stationr	6.00
chessboard/checkerboard	SlabThn2	SportGam	3.50
cracker	SlabThn2	Food	1.50
cupboard door	SlabThn2	FurnGen	4.00
door	SlabThn2	BldgPart	5.50
DVD box	SlabThn2	ElctrnAV	2.75
DVD case	SlabThn2	ElctrnAV	2.75
exercise book	SlabThn2	Stationr	3.25
floor tile	SlabThn2	BldgMatr	3.25
floppy disk 3.5"	SlabThn2	ElctrCmpt	2.00
game cassette	SlabThn2	ElctrCmpt	2.00
gum pack (Excel)	SlabThn2	FoodCont	2.00
gyprock (drywall)	SlabThn2	BldgMatr	6.00
mini disk	SlabThn2	ElctrCmpt	2.00
mousepad	SlabThn2	ElctrCmpt	2.75
pop tart	SlabThn2	Food	2.00
school desk (top)	SlabThn2	Stationr	3.75
tile	SlabThn2	BldgMatr	2.75
tooth	SlabThn2	Animal	1.00
bleachers	Stair	SportGam	8.00
steps/staircase	Stair	BldgPart	6.00
air hockey table	TableStr	SportGam	5.25
bed frame	TableStr	FurnBdrm	5.50
coffee table	TableStr	FurnGen	4.75
computer desk	TableStr	FurnGen	4.75
desk1	TableStr	FurnGen	4.75
table	TableStr	FurnGen	5.50
tool bench	TableStr	Tools	5.50
billiard rack	Triangle	SportGam	3.25
guitar pick	Triangle	Music	1.00
shark tooth	Triangle	Animal	1.00
triangle	Triangle	Shape	-----
yield sign	Triangle	StrtSign	4.00
baseball dugout	basbldug	SportGam	7.00

blade of grass	bldgrass	bldgrass	1.00
carpenter's square	carpntsq	Tools	3.25
chair	chair	FurnGen	4.50
comb	comb	PersHygn	2.00
couch	couch	FurnGen	5.50
car chassis	crchassis	Vehcl	6.75
cross	cross	Symbolic	-----
door frame	doorfram	BldgPart	5.50
Empire State Building	empStBld	Bldgs	10.00
hockey stick	hockeyst	SportGam	4.00
medicine bottle	medbottl	DrugMed	2.00
clock tower at MUN	muclockt	BldgPubl	7.00
pair of pants	parpants	ClothngR	4.25
pinball machine	pbllmach	SportGam	5.50
A-frame house	prism	Bldgs	9.00
pyramid	pyramid	BldgPyrm	-----
radiator grill	radgrill	Vehcl	4.50
shopping cart	shpgcart	shpgcart	4.75
skateboard	sktboard	SportGam	3.75
staple	staple	Stationr	1.00
stop sign	stopsign	StrtSign	4.00
paddle (straight-sided)	strpaddl	SportGam	4.25
treadmill	tredmill	SportGam	5.50
doorstop	wedge	BldgPart	1.75