VISUAL OBJECT REPRESENTATIONS:
EFFECTS OF FEATURE FREQUENCY AND SIMILARITY

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

SELDA EREN KANAT

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
THE DEPARTMENT OF COGNITIVE SCIENCE

DECEMBER 2011

Approval of the Graduate School of Informatics

Prof. Dr. Nazife BAYKAL

Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Doctor of Philosophy.

Prof.Dr.Deniz ZEYREK

Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Doctor of Philosophy.

Assist. Prof. Dr. Annette HOHENBERGER

Supervisor

Examining Committee Members

Prof. Dr. Deniz ZEYREK	(METU, COGS)	________________
Assist. Prof. Dr. Annette HOHENBERGER	(METU, COGS)	________________
Assist. Prof. Dr. Didem GÖKÇAY	(METU, COGS)	________________
Assist. Prof. Dr. Mine MISIRLISOY	(METU, PSY)	________________
Assist. Prof. Dr. Emre ÖZGEN	(BİLKENT, PSY)	________________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this wok.

Name, Last name: Selda Eren Kanat

Signature : _________________

ABSTRACT

VISUAL OBJECT REPRESENTATIONS:
EFFECTS OF FEATURE FREQUENCY AND SIMILARITY

Eren Kanat, Selda

Ph.D., Department of Cognitive Science

Supervisor: Assist. Prof. Dr. Annette Hohenberger

December 2011, 185 Pages

The effects of feature frequency and similarity on object recognition have been examined through behavioral experiments, and a model of the formation of visual object representations and old/new recognition has been proposed. A number of experiments were conducted to test the hypothesis that frequency and similarity of object features affect the old/new responses to test stimuli in a later recognition task. In the first experiment, when the feature frequencies are controlled, there was a significant increase in the percentage of “old” responses for unstudied objects as the number of frequently repeated features (FRFs) on the object increased. In the second experiment, where all features had equal frequency, similarity of test objects did not affect old/new responses. An evaluation of the models on object recognition and categorization with respect to the experimental results showed that these models can only partially explain experimental results. A comprehensive model for the formation of visual object representations and old/new recognition, called CDZ-VIS, developed on the Convergence-Divergence Zone framework by Damasio (1989), has been proposed. According to this framework, co-occurring object features converge to upper layer units in the hierarchical representation which act as binding units. As more objects are displayed, frequent object features cause grouping of these binding units which converge to upper binding units. The performance of the CDZ-VIS model on the feature frequency and similarity experiments of the present study was shown to be closer to the performance of the human participants, compared to the performance of two models from the categorization literature.

Keywords: Visual Object Representation, Feature Frequency, Discrete Feature Similarity, Old/New Recognition, Convergence Divergence Zone Framework

ÖZ

GÖRSEL NESNE GÖSTERIMLERI:
ÖZNITELIK FREKANSININ VE BENZERLİĞİN ETKISI

Eren Kanat, Selda
Doktora, Bilişsel Bilimler Bölümü
Tez Yöneticisi: Yrd. Doç. Dr. Annette Hohenberger

Aralık 2011, 185 Sayfa

Bu tezde, öznitelik frekansı ve benzerliğinin nesne tanıma üzerine etkisini inceleyen bir dizi deney yapılmış, görsel nesne gösterimlerinin oluşumu ve eski/yeni tanımayı açıklayan bir model önerilmiştir. Öncelikle bir dizi deney gerçekleştirilmiş, bu deneylerde, katılımcılara gösterilen nesnelerin özniteliklerinin frekans ve benzerliklerinin nesne tanımlama safhasındaki eski/yeni yanıtlarını etkilediği hipotezi test edilmiştir. Birinci deneyde, öznitelik frekansları kontrol edildiğinde, yeni nesneler üzerindeki "sık tekrar eden öznitelik" sayısı arttıkça bu nesneler için verilen "eski" yanıtlarının oranının arttığı tespit edilmiştir. İkinci deneyde, tüm öznitelik frekansları eşit tutulduğunda, test nesnelerinin çalışma nesnelerine benzerliğinin eski/yeni yanıtlarına etkisi olmamış, yüksek oranda "eski" yanıtı verilmiştir.

Bu sonuçlar "sahte anı" literatüründeki kategori etkisi ile ilgili bulgular tarafından desteklenmektedir. Nesne tanıma ve sınıflandırma modelleri deneysel sonuçlar üzerinden incelendiğinde bu modellerin deneysel sonuçları yalnızca kısmen açıklayabildiği görülmüştür. Bu nedenle Yakınsama-Iraksama Bölgesi Platformu (Damasio, 1989) üzerinde geliştirilmiş olan CDZ-VIS modeli önerilmiştir. Yakınsama-Iraksama Bölgesi Platformu'nun ana prensibi birarada görülen nesne özniteliklerinin hiyerarşik nesne gösteriminin üst katmanlarında bir bölgeye yakınsamasıdır. Nesneler gösterilmeye devam ettikçe sık tekrar eden öznitelikler bu bağlantı bölgelerinin gruplaşmasına ve daha üst katmanlarda bir bölgeye yakınsamalarına yol açar. Son olarak, CDZ-VIS modelinin bu çalışmada gerçekleştirilmiş olan öznitelik frekansı ve benzerliği deneylerinde gösterdiği performansın, sınıflandırma literatüründeki iki modelle kıyaslandığında katılımcıların performansına daha yakın olduğu gösterilmiştir.

Anahtar Kelimeler: Görsel Nesne Gösterimi, Öznitelik Sıklığı, Ayrık Öznitelik Benzerliği, Eski/Yeni Tanıma, Yakınsama - Iraksama Bölgesi Platformu

This work is dedicated to;

Barbun – The soulmate of Tekir

&

My family

ACKNOWLEDGEMENTS

I would like to express my sincerest appreciation to my advisor Assist. Prof. Dr. Annette Hohenberger for her kind supervision, guidance, and friendly attitude throughout this study. I would like to thank Assist. Prof. Dr. Didem Gökçay and Assist. Prof. Dr. Emre Özgen for their precious guidance and Assist. Prof. Dr. Mine Mısırlısoy and Assist. Prof. Dr. Murat Perit Çakır for their valuable comments and suggestions about this study. Special thanks to Prof. Dr. Nazife Baykal for her support and encouragement throughout my studies at the Informatics Institute.

I would like to express my thanks to Demet Kırbulut, Sevda Altay and Özge Alaçam for being my best friends and sharing all my happy and sad days, and for encouraging their colleagues to participate in the behavioral experiments of this study. I would also like to thank Necla Işıklar and Sibel Gülnar for their help during the official procedures.

Special thanks are extended to my parents Selmin Eren, Ataullah Eren, Gülay Kanat, Muhittin Kanat and my sisters Eda Erkan, Necibe Emel Kanat for being by my side during my studies.

Finally, I would like to thank my husband İrfan Emrah Kanat, for showing patience to my endless brainstorming and hypothesizing. He is the only person who was exposed to all the consequences of the difficulties I have experienced during the process.

I acknowledge The Scientific and Technological Research Council of Turkey (TUBITAK) for supporting this work through the National Scholarship Program with a PhD scholarship.

TABLE OF CONTENTS

LIST OF TABLES

TABLE

LIST OF FIGURES

FIGURE

LIST OF ABBREVIATIONS

ANOVA: Analysis of Variance

CDZ: Convergence Divergence Zone

CDZ-VIS: Convergence Divergence Zone Visual

DS: Discrete Similarity

FRF: Frequently Repeated Features

GCM: Generalized Context Model

HF: Hippocampal Formation

LTM: Long Term Memory

MDS: Maximum Discrete Similarity

PHC: Parahippocampal Cortex

PHR: Parahippocampal region

PRC: Perirhinal Cortex

STM: Short Term Memory

RT: Reaction Time

TEC: Theory of Event Coding

CHAPTER 1

INTRODUCTION

Evolution has led to neural systems that can process specific input and display a particular behavior. It is interesting that by simple mechanisms of input processing these systems can construct representations of their environment which in turn affects how the new input will be processed. Combined with a body that can act on its environment, these systems gain access to a vast amount of input sources and even modify these sources.

For a researcher determined to understand the formation of representations and how they affect the system in turn, it seems straightforward to try to understand these processes by examining simple processes executed by neurons. Single cell recording is thus the most direct way of providing explanations for how these small units of processing communicate with each other to form representations of the environment. But as a gift of evolution, these small units are capable of interacting in various ways, and displaying different behaviors under different conditions and among different assemblies. Considering the number of neurons in the brain, it seems infeasible to examine the neurons under every condition. This led to simulations of neurons that enabled automated testing of these conditions.

From a higher point of view, we see cell assemblies which have their own dynamics. Since single cell recording is the only measure of direct activity, it is not possible to directly examine the behavior of neural populations. The artificial neural network literature contributed to this level of analysis. Competitive, cooperative, self-organizing networks were shown to replicate the behavior of neural populations for various cases. Explanations of behavioral observations with population dynamics were interesting since they included both the neural and behavioral phenomena (Gerstner, 2000; Matsumoto, Okada, Sugase-Miyamoto, Yamane, & Kawano, 2005; Usher, Schuster, & Niebur, 1993).

The oldest way of examining the capabilities of the neural systems is through observations of the behavior. The downside of this approach is that the interactive nature of the neural systems leads to a space of infinitely many behaviors. A reliable research requires elaborate planning of the experiments: controlling the variables to measure what is really intended to be measured. Years of research led to standardized methods of experimentation and analysis of the results.

The methods for providing explanations of the formation of representations except experimental psychology are relatively new and still developing. Thus, the number of papers published by these methods is still low. What we know about the formation of representations comes mainly from the experimental psychology literature. Recently, the computational cognitive neuroscience literature has offered some models of the formation of visual object representations (Rolls & Deco, 2002), but they are specific to a single phenomenon, like viewpoint invariance, and they do not form a basis for explaining behavioral findings.

This study adopts two methods: a behavioral experiment and a model analysis. The aim of the behavioral experiment is to show the effect of two important factors on the formation of visual object representations: feature repetition frequency and discrete similarity. Despite recent evidence regarding the feature-based structure of object representations, the number of experimental studies which examine feature induced effects on memory is quite low. The model analysis in this study involves the formulation of a model of the formation and structure of visual object memory according to the recent literature and results of the behavioral experiments conducted for this study. The model has been implemented and tested with stimuli from the behavioral experiment, and compared with human data.

1.1 Problem statement

This study investigates the following research question:

- How do representations of objects and categories emerge from object features?

The solution to this problem requires a statement of the specifications of a mechanism for the formation of visual object representations and categories. A mechanism can be specified by either an explanation of its principles or building a working model of the mechanism. There are many advantages of building a working model over giving only a descriptiption of principles (Dawson, 2004; McClelland, 2009). First, a model is testable, by using experimental data as the validation source. Second, a model has predictive power, which results in new hypotheses to be tested experimentally or new phenomena imposing new research questions.

In the first part of this dissertation, behavioral experiments were performed to test a number of hypotheses about the effect of feature frequency and similarity on object recognition. In the second part, the results of the experiments and findings from cognitive psychology and cognitive neuroscience were used to develop new hypotheses about the possible mechanism responsible for the formation and structure of visual object representations.

1.2 Significance of the study

Visual memory for objects is required for perceiving, manipulating and communicating about objects. Visual LTM has close connections to the areas of visual perception, action and language. Each of these areas has its own literature. However, there is no unified theory which can explain the following interrelated aspects of visual memory:

- The structure of object representations
- The formation of object representations
- The effect of object representations on perception of new and previously perceived objects
- The link between object representations and possible actions on them (affordances)
- The link between object representations and categories
- The link between object representations and language

The importance of a theory which explains the above capabilities is that it would lead to a comprehensive and unified explanation of various cognitive phenomena. An interesting thing about the existing memory literature on visual memory is that there is almost no effort to construct such a theory. Instead, there have been hundreds of micro-studies that examined the effect of some controlled variables to test a specific hypothesis. Of course, in order to construct a comprehensive explanation, there should be enough evidence for constructing the building blocks of the theory. However, the hypotheses tested by micro-studies cover only such a tiny portion of the actual phenomena that it is almost impossible to combine them to come up with a unified theory.

In this study, a mechanism for the formation of visual object representations is proposed. First, the solution to the problem was hypothesized, and experiments were conducted to test hypotheses related to this solution. After reviewing a vast amount of existing models with respect to the experimental findings of the present study, the plausible and implausible components were identified and using the experimental data and findings from neuroscience and cognitive psychology as guides, a new model, called CDZ-VIS, was developed. This model is an attempt towards a comprehensive and unified explanation of the phenomena related to object recognition, whose aspects are listed above.

1.3 Organization of the Dissertation

There are seven chapters in this dissertation. In the first chapter, that is, the current chapter, the subject of the study has been introduced and the problem addressed by the study has been stated. In the second chapter, a literature survey is presented which explains current findings relevant to the solution of the problem.

The third chapter starts with the statement of the hypotheses of the study. The methods for testing these hypotheses and the design of the experiments are explained, and the results are presented. A discussion of the results is provided at the end of the chapter.

The fourth chapter presents a detailed analysis of existing models of object recognition and memory, in terms of the experimental results stated in the third chapter. The solutions offered by these models to the problem of visual object representations are evaluated, and shortcomings are identified.

In the fifth chapter, the proposed model of the structure and formation of visual object representations is presented. The framework it was built on and formal specifications of the model are explained.

In the sixth chapter, a quantitative and qualitative comparison of the proposed model with the related models in the literature is provided. In the quantitative comparisons, two prominent models were selected and implemented for comparison. For the qualitative comparison, models in the false memory literature were discussed in comparison to the principles of the CDZ-VIS model, including a separate discussion for the connectionist models of memory.

The last chapter summarizes the problem addressed in this study, and the solution offered to this problem. The results of the experiments and modeling study are briefly stated, and the implications of the proposed model for the literature on visual object representations are discussed. The chapter ends with stating some limitations of the study and suggestions for future studies.

CHAPTER 2

LITERATURE REVIEW

This section aims to present findings from studies on the structure of object representations in memory, from behavioral and neurobiological research. First, findings on the structure of object representations are reported, and information on different aspects of these representations such as invariance, intensity and constitution are given. Second, neural correlates of object representation in the human brain are provided according to recent findings in the domain. Third, the top-down effects of object representations on perception and mental imagery are discussed. Finally, recent literature on the close link between perception and action is provided.

Findings from object recognition and categorization studies are also presented since they make implicit assumptions about the structure of visual object representations and thus provide valuable clues about the memory representations of objects.

2.1 The visual LTM

Evidence from current studies is accumulating that converges on a common understanding of the nature of the visual representations (Tyler, Likova, & Nicholas, 2009). According to this evidence, distributed representations of low level features in the primary visual cortex bind together to construct high-level object representations in more anterior parts of the brain. The details about these representations are explained in this section.

2.1.1 Objects, Categories and Concepts

The term 'visual object representation' has been interpreted differently by different disciplines of Cognitive Science. In the object perception literature, visual object representations are considered as the representation of shapes and patterns. Higher levels of representations are regarded as top-down effects in object recognition. In the categorization literature, visual object representations correspond to representation of object categories rather than individual objects. Low-level feature processes are kept out of the scope of interest. In the memory literature, object representations are considered as sensory memory, in a similar way to the object recognition literature. Individual representations of objects are not considered as a visual representation, but classified under the declarative memory as conceptual representations of objects. It is important to keep in mind that different disciplines define the problem in their own terminology and point of view. The following review of the literature therefore includes findings from neuroscience to psychology, which is essential to provide a unified explanation of the phenomena related to visual object representations.

2.1.2 Neurobiological correlates

The first step of formation of visual memories is extracting the information from visible light. As soon as the signals are sent from the retina to the brain, the processing has begun. Different kinds of information are transferred through the optic nerve, such as different wavelengths, temporal dynamics caused by the motion of the stimulus, spatial distributions, etc. Each type of information travels through the brain towards different regions where they are processed for various purposes. Although the number of paths is large, they can be subsumed under two main paths. The first path is a primitive one, also found in invertebrates. It goes from the optic tectum of the midbrain, superior colliculus connected to premotor and motor nuclei through the thalamic nuclei to the visual sites in cerebral cortex. The second path is found in primates and goes towards the thalamus, to the dorsal part of lateral geniculate nucleus, then to cerebral cortex and finally to the primary visual area, also called V1 or striate cortex in the occipital lobe.

Dorsal/ventral distinction

In the human vision literature, there are two well-known pathways from the primary visual area: a ventral one leading to the inferior temporal lobe, and a dorsal one leading to the posterior parietal lobe (Mishkin & Ungerleider, 1982). The ventral path is thought to transmit object properties and the dorsal path spatial properties. Although this is a commonly accepted view, there are alternative ideas which claim that the dorsal path is involved in calculating visuomotor transformations for the purposes of motor actions. On the other hand it is claimed that the ventral path is used for perception and cognitive manipulation. Since object properties like shape and size are also important in visuomotor transformations, the ventral path should also support the processes in the dorsal stream. The calculations of transformations incorporating shape and size information are performed in the dorsal system (Dijkerman, Milner, & Carey, 1996).

Role of sub-cortical structures

There is strong evidence for the role of the parahippocampal region in storing visual long term memories. This region consists of perirhinal cortex (PRC), entorhinal cortex (EC), and parahippocampal cortex (PHC). In the right hemisphere, parahippocampal gyrus was found to be related to delayed recall (Kohler et al., 1998). Kohler et al. (1998) found an association between the loss of parahippocampal gyrus and delayed non-verbal visual memory deficits in Alzheimer's disease.

In addition to patient studies, the use of new imaging techniques enabled researchers to study regions of the brain that are responsible for specific types of memory. Duzel et al. (2003) examined the regions related with associative memory in a task where subjects had to remember associations between a face and a tool presented together with that face. The associations depended on the spatial arrangement of the tool relative to the face or identity of the tool. Recordings of cerebral blood flow showed that when the association presented in the test phase was recognized, PHC was highly active. Conversely, when the presented association was novel both in terms of the spatial arrangement of the tool and the identity of the tool, PHC was less active. The region which was activated during the presentation of novel associations was the hippocampal formation (HF), which was also more active for the changes in spatial position of the tool with respect to the face, rather than the identity of the tool itself. The associations that were presented in the familiarization phase did not cause much activation in HF during the test phase. These results indicate that PHC has a role in visual long term memory whereas HF is related with identification of novel stimuli.

The role of PHR and HF in recognition of previously presented stimuli has also been studied by Ranganath & D'Esposito (2005). They examined recognition in terms of two sub-processes: recollection in context and familiarity with the stimulus features. Taking this division as a basis of their analysis, they found that the main contribution of the hippocampus is to recollection. In familiarity tasks, the hippocampus did not display much activation compared to the activation during the recollection tasks. Additional activation was found in PHC which might indicate a relation between PHC and recollection. Ranganath & D'Esposito assume that this contribution is possible since PHC can represent and retrieve contextual information. The form of this contextual information is mostly spatial. Thus, hippocampus and PHC together support recollection in context. For the familiarity part, they could show that PRC was active during familiarity tasks and they conclude that PRC is necessary for recognition by familiarity.

Localization studies

Objects are represented by a set of features distributed throughout the brain. These features are organized hierarchically, from lower levels to higher levels. The visual ventral stream is shown to be the primary location for object representations. However, several regions contribute to these representations. The modularity of the visual system has been challenged by various fMRI studies (Bussey & Saksida, 2005). Evidence from fMRI studies indicates that perirhinal cortex encoding individual object information and posterior parahippocampal cortex encoding context information connect to the hippocampus for the integration (Davachi, 2006). Perirhinal cortex plays a role also in familiarity decisions (Eichenbaum, Yonelinas, & Ranganath, 2007).

"Mind reading" studies provide important evidence for localization of object representations. These studies aim to predict the identity and/or the category of the objects viewed by the participants from fMR images. Shinkareva et al. (2008) report that particular regions allow good predictions of object identity: the bilateral SES, IES, calcarine sulcus, fusiform gyrus, IPS, left IPL, posterior superior, middle and inferior temporal gyri, postcentral gyrus,

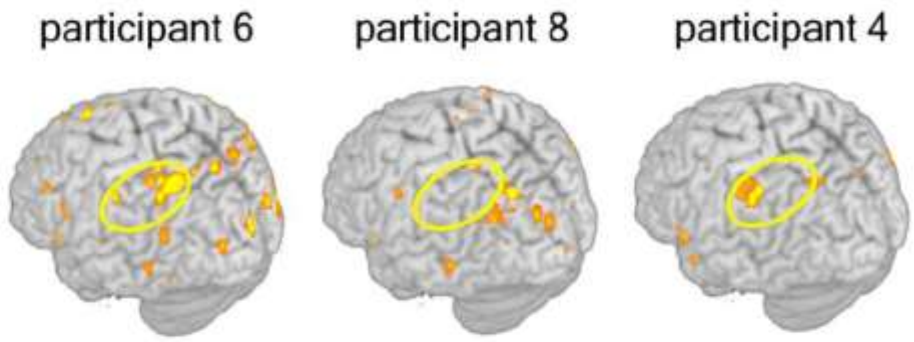


Figure 1. Object exemplar classification. The region marked with yellow ellipses indicates the common voxels used for classification. Three participants having highest accuracies are shown (Shinkareva et al., 2008).

and hippocampus (Figure 1). Tools caused activation mainly in the left hemisphere, especially in the ventral premotor cortex and posterior parietal cortex. For dwellings, right parahippocampal gyrus (close to parahippocampal place area) was useful for prediction (Figure 2-B). For predicting object categories, bilateral SES, calcarine, IES, SPL, IPL, IPS, fusiform, posterior superior and middle temporal, posterior inferior temporal gyri, cerebellum, and left precentral, superior frontal, inferior frontal triangularis, insula, and postcentral gyri proved to be useful (Figure 2-A). Within-subject predictions were more successful than between-subjects predictions. This shows that individual variations play an important role in localization. Another result they obtained is that it was possible to make predictions using individual regions. Even though brain activation was distributed across the whole brain, predictions from a single region were also successful. This might indicate that different parts of the brain encode different properties of the same object. For example, one area may represent the possible hand position to manipulate the object whereas another area might be sensitive to the shape of the object.

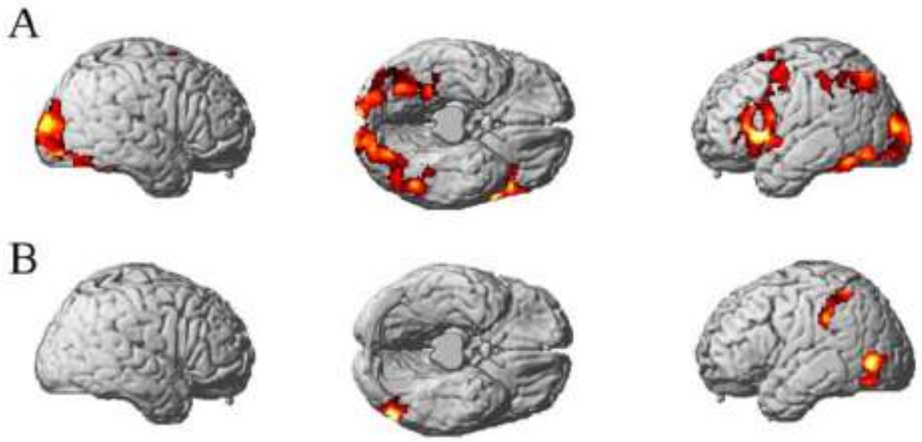


Figure 2. Results from object localization studies (A) Objects compared to fixation (B) Tools compared to dwellings (Shinkareva et al. 2008).

Another research field related to the localization of object representations is object recognition. Low-level object areas have been studied well whereas location of higher level representations of individual objects is still unclear. The reason for such discrimination might be that the activations in low-level visual areas like V1 and V2 could be linked directly to the incoming stimuli but higher level activations may not have such direct connections with the outside world; they can be any abstract entity derived from the features of the stimuli according to the needs of the organism. Besides, high level representations activate many regions in the whole brain simultaneously, whose functions are currently unclear. However, studies on object recognition are providing some clues on these representations. Relevant to the experiments presented in this dissertation, there are object familiarity studies which explore the neural basis of false memories. Danker & Anderson (2010) presents a good review of such studies. They conclude that the level of activity is highest for hits, moderate for false alarms / false negatives and lowest for correct rejections. Wheeler & Buckner (2003) show that these activations take place in the left parietal cortex near BA 40/39. Despite the vast amount of studies on the localization of object representations, what features are encoded in the detected regions, how they are encoded, how the low-level features bind and activate individual object representations and abstract categories are poorly known (Thompson-Schill, 2003).

2.2 Structure of visual object representations

2.2.1 Structure

A set of features distributed in different cortical regions constructs visual LTM, and these features bind together to form coherent representations of objects (Slotnick & Schacter, 2004). Object representations are stored non-topographically in visual LTM. These representations contain both visual and spatial features as well as features in other modalities like the auditory and tactile modality. Associated features are connected to each other so that the related features can be retrieved whenever needed. These features are integrated in visual STM in this case (Luck, Girelli, McDermott, & Ford, 1997).

Several studies determine the locations of different types of features. The most studied types are visual and spatial features. Neuroimaging studies reveal that during tasks of object identity retrieval, inferior temporal cortex in the fusiform gyrus is activated. The region of activation is different in spatial location retrieval tasks, in which the inferior parietal lobe in the supramarginal gyrus is activated (Kohler et al., 1998). Similarly experiments were conducted to determine whether color and shape information are represented in different regions of LTM. Results confirmed the hypothesis that color is one of the features which constitute the object representation as well as the shape of the object. It was also shown that these features are stored and retrieved separately (Hanna & Remington, 1996).

Spatial properties are the relative positions of two or more objects or object parts (Kosslyn, 2005). Spatial representations can be divided into two categories according to the frame of reference in which they are formed, namely subjective and objective. A subjective representation is relative to the observer whereas an objective representation is relative to the environmental coordinates. The basic form of spatial representations in humans is thought to be subjective. It is easier to form since the point of reference is the observer him/herself, which is constant over the course of observation. However, it is not sufficient in the everyday world since humans also need objective representations when they need to locate objects with respect to other objects in the environment. Studies with infants showed that humans gain the ability to form and use objective spatial representations at 4 months of age (Bremner, Bryant, Mareschal, & Volein, 2007). From this age on, humans can form and manipulate object-centered spatial configurations to accomplish complex tasks in the environment involving multiple objects.

In a set of experiments, Xu (2002) presented participants a set of images during the familiarization phase of an object recognition task. In the recognition phase subjects were asked to recall features belonging to the images. The results showed that it was much easier for the subjects to remember features that belonged to the same object. When they were asked to recall features belonging to two different objects, subjects were less successful. This was true also for the parts of an object. When the stimulus was an object consisting of distinct parts, subjects were much better in recognizing the features from the same part of the object, rather than features from different parts of the object. Kahneman, Treisman, and Gibbs (1992) argue that visual information is stored in ‘object files’, instead of ‘place files’. Thus, visual attention is directed at objects.

2.2.2 Feature-based object representations

For the past two decades, feature-based object representations are favored in both computer and human vision studies. The main reason for this preference lies in the feasibility of dealing with real world objects by means of such representations. Previously, researchers held the idea that objects are represented individually in the form of a sketch or depiction. The visual system was thought to perform like a camera in that it receives the light through the eyes, whereupon the light leaves traces in the brain and these traces make up the visual memories. However, such representations did not seem to be suitable for processing, since the same object is hard to recognize under different illumination conditions, from different perspectives, or being (partially) occluded by other objects. Perspective- and illumination-invariant representations have been favored, maintaining only the features which are relevant to identifying that object.

Another problem arises at this point: Which features should be extracted? When the visual flux first arrives at the retina, it consists of a large number of dimensions. It is a challenge for the visual system to extract a useful set of dimensions from this flux in order to construct efficient representations (Cutzu & Edelman, 1996). If the feature space is not restricted by the processes that construct the object representations in LTM, the manipulation and search through these representations would have been computationally intractable. This is a common problem in the domain of computer vision, where the sensors receive a continuous amount of visual information from the environment. In tasks such as object recognition and object manipulation efficient representations of these objects are required. Thus, reducing the number of features extracted from the scene by methods like feature selection, principle component analysis, etc. have been preferred. However, current studies on object recognition in humans and primates indicate that the visual system may also store view-point specific information about objects (Epstein, Graham, & Downing, 2003).

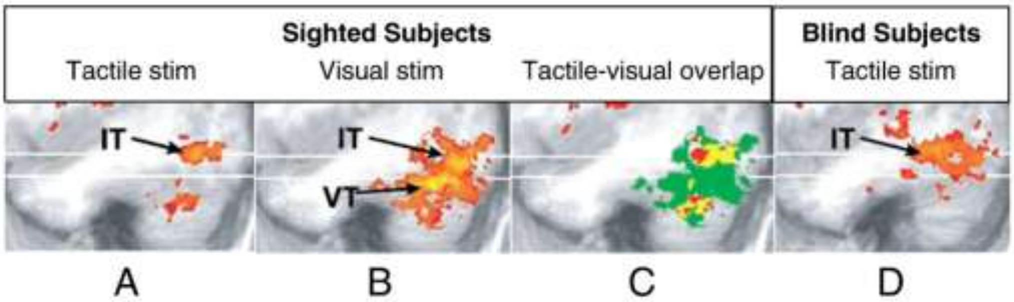


Figure 3. Features of the same object in different modalities overlap in IT. A) Activated region when the tactile stimulus is presented B) Activated region when the visual stimulus is displayed C) The overlap between activations of tactile and visual stimuli D) The activated region when the tactile stimulus is presented to the blind subjects. Adapted from (Tyler et al., 2009).

2.2.3 Intensity of representations

Object representations in visual LTM may have different intensities. The graded nature of these intensities shows its dominance in object recognition tasks, where object-based effects are tested (Ariga, Yokosawa, & Ogawa, 2007). In one task, subjects were asked to recognize a target object in different conditions. In the first condition, the object was presented with a cue and in the second condition with no cue. Subjects were faster at responding to objects presented with a cue only when the displayed object has an LTM representation of high intensity.

The main factor determining the intensity of an object representation in LTM is the amount of attention directed to that object. Even when the stimulus is noisy or low quality, attending to the object increases the strength of its representation in LTM. Conversely, although an object might be displayed at a high quality format, its representation will not be of high intensity if the subject does not pay particular attention to that object. Thus attention is crucial for consolidating adequate object representations in LTM.

Two effects of attention on object processing have been claimed: Sensory enhancement and prioritization (Lee & Vecera, 2005). Sensory enhancement corresponds to the improvement of object representations by attention. Prioritization is the change in the order of analysis of the scene according to attention processes. It provides an advantage for the identification of the attended objects since the parts of the object that were unattended in the first look will be attended in the next one (Ariga et al., 2007).

2.3 Object recognition

2.3.1 From V1 to IT

When visual stimuli arrive at V1, they are represented in a topographically organized manner. The object properties follow the ventral path and the spatial properties the dorsal path. These two types of information arrive at the sites of visual LTM where object representations reside (see Figure 3). Combined with other information like the smell of the object, a particular representation can be activated there. There are top-down connections from visual LTM to the topographically organized primary visual cortex. According to Kosslyn (2005), the features that constitute the activated representation are projected backwards onto V1 via these connections. The attention mechanism searches for these features in the scene. The new stimulus arrives at V1 and travels through the dorsal and ventral pathways towards the visual LTM, and activates an object representation. If the activated representation is the same as the one in the previous cycle, its activation is strengthened. If it is a different one, although it might share some features with the previous one, new features are projected onto V1, directing attention to those features in the scene. This process continues until a single object representation is determined in LTM. Visual memory and visual perception are associated with common neural substrates (Slotnick & Schacter, 2004).

The finding that stimuli from the scene activate object representations in LTM has been confirmed in other studies. According to these studies, orienting attention changes activity levels in LTM structures. These structures are mainly the fusiform and parahippocampal gyri (Summerfield, Lepsien, Gitelman, Mesulam, & Nobre, 2006).

2.3.2 Top-down effects

The emphasis on top-down effects in perception is changing through the decades in the history of the perception literature. When researchers discovered top-down effects, they started to ignore the importance of bottom-up processes. In the 1980s, bottom-up processes were supported against the over-interest in top-down processes. Thus, Marr (1982) claimed that objects stimulate relevant brain regions when they come into the scene and match with the corresponding object representations in LTM. By the 1990s top-down processes were postulated behind seemingly pure bottom-up processes. Evidence from priming studies showed that representations in LTM that are active before a recognition task cause familiar objects to be remembered better (Li & Yeh, 2007), reminding of the importance of top-down processes. This effect is so specific that when the activated representations are in the same modality with the presented objects, the probability of recall is much higher. Li and Yeh (2007) shows that novel objects have no special property of gaining attention, and rejected a stimulus-driven account of perception. Recently, a more balanced view is dominant. Perception is now seen as a two-way process, with bottom-up and top-down processes interacting with each other (Kosslyn, 2005).

2.3.3 Familiarity

Verde and Rotello (2003) define familiarity as “the nonspecific sense of ‘oldness’ produced when an object matches the contents of memory” (p. 739). From a different point of view, Yonelinas (1997) defines familiarity as a signal detection process.

Familiarity studies consist of presenting a number of items to the participants first and then asking the familiarity of another set of items. Several factors were shown to affect the familiarity of these items. For example, according to Azimian-Faridani and Wilding (2004), intervening tasks presented before the old/new decision decrease familiarity in general.

Watkins and Peynircioglu (1990) coined the term revelation effect to describe the increased tendency to call an item “old” when it is not displayed directly during the test phase but revealed, or discovered, by the participant. Anagrams can be example of such test stimuli: subjects try to find the word hidden in the anagram. Revealing the test item slowly affects familiarity (Verde & Rotello, 2003). Mulligan and Lozito (2006) state that the revelation effect is usually larger for the new items.

Landau (2001) reports studies which define the revelation effect as a problem of source attribution. If before the recognition decision, the source of familiarity is attributed to the study list, instead of the intervening task. However, later research showed that the revelation effect occurs even when the practiced item is different from the test item. For example, solving an anagram of APPLE increased the familiarity for CHAIR. This is in conflict with the source attribution explanation.

Table 1. Existing models for explaining false memory effects.

Theory	**Explains**	**Cannot explain**	**Type**
Implicit Associative Response (Underwood, 1965)	probability of false recognition increases as the number of associates seen before increases	False memory for artificial items	Encoding
PDP (Rumelhart & McClelland, 1988)	Associative effects	Typicality effect	Encoding Retrieval
Fuzzy trace (Brainerd & Kingma, 1984)	Word superiority effect; change in hits in time	Effects of item frequencies	Retrieval
Prototype theory (Rosch, 1975)	Typicality effect, output dominance	Associative effects	Encoding Retrieval
Activation monitoring (Roediger, Watson, McDermott, & Gallo, 2001)	probability of false recognition increases as the number of associates seen before increases	False memory for artificial items	Encoding/ retrieval

Table 1 (cont.)

Source Monitoring (Johnson, 1988)	Effects resulting from information from other sources	False memory for artificial items	Retrieval
Attributional model of memory (Kelley & Jacoby, 1990)	When fluency increases, false recognition increases	How related items lead to false recognition of critical items	Decision-based
Illusory conjunctions (Kroll, Knight, Metcalfe, Wolf, & Tulving, 1996)	Binding problems	Effects of item frequencies	Encoding

2.3.4 False memory

Object recognition does not always result in an accurate identification of the presented object. An object can be falsely identified as new even if it is has been presented before. On the other hand, an unstudied object can be recognized as studied. These erroneous responses are in the scope of the false memory research. Even though the name implies a "memory-based" explanation, research in this area involves processes other than memory, like decision making processes. Also, the memory processes are not taken as a single mechanism: encoding and retrieval processes are considered separately.

Thus, accounts of false memory phenomena can be examined under two groups: memory-based and decision-based. A list of these accounts and what can and cannot be explained by these accounts can be seen in Table 1. Proponents of the memory based accounts claim that false memory occurs because of encoding and retrieval of presented items. Decision-based accounts attribute false memory effects to some errors during the decision making process, rather than item representations. Both accounts have their own supporting evidence, and there is ongoing debate about the plausibility of each. Existing explanations (models/theories) in the current false memory literature can be seen in Table 1, including the effects that they can and cannot explain.

The prototype theory by Rosch (1975) has been influential on memory-based accounts of false memory, especially for explaining category-induced effects on false recognition. It explains the category structure as graded where each member of a category has varying degrees of category membership. This graded structure is reflected in experiments where people list the members of a category. Some category members are listed more frequently than others, and this frequency is called category output dominance. In addition, people can give ratings for how typical each member is of the specified category, which is called the typicality of the member. Typicality is seen as a function of similarity of members to the category prototype (Schmidt, 1996). Smith, Ward, Tindell, Sifonis, and Wilkenfeld (2000) conducted an experiment to test the effects of typicality and category output dominance on the false recognition of category members. They hypothesized that more typical members of a category would be more often falsely recalled than less typical members of the category, and claimed that the reason for this is the easier access to the typical members, since he assumes that recall is related to the similarity of items to the general idea of an experience (which can be a category). Following this line of thought, they discuss that there can be other measures of graded category structure which are more related to the easier access concept, and suggests using category output dominance.

Explanations by Smith et al. (2000) are memory-based, with reference to the prototype theory, namely they claim that false memories are caused by the graded stucture of categorical representations. Graded category structure is a claim of the prototype theory. They list a number of measures that indicate graded categories: typicality, output dominance, central tendency, and category ideals. In their experiments, they compare the effects of typicality and output dominance on recall. Output dominance is the frequency of an item when items of a category are to be listed (Barsalou, 1985) whereas typicality is the subjective rating of the participants for how typical the item is of a category.

Memory-based accounts are not restricted to category effects. In their earlier paper, Roediger and McDermott (1995) supported associative response models. According to the associative response account, an item shown during the study phase activates memory representations of associated items, and residual activation from these representations cause unstudied items to be remembered as old. However, this explanation cannot account for the results of the present study. The items were created artificially, thus an unstudied test item does not have a previously built memory representation. Even though the individual features are familiar, their combinations are novel. As a result, any model that claims an activation of associated responses during the study phase is not capable of explaining the false memory effects for artificially created stimuli.

2.3.5 Contextual effects

There is evidence for both context-dependent and context-independent representation of objects. The classical color adaptation phenomena indicate that colors can be perceived differently when the neighbor colors change. However, this is not a direct evidence for a context-dependent representation of color. It is possible that the individual representations of colors are context-independent but different representations might become active when neighboring colors change (Tyler et al., 2009).

2.3.6 Missing features

It is important to consider the case when the incoming information is incomplete. The visual system is flexible and adaptive so that the organism can survive in dynamic environments. An object should be identified as fast and correctly as possible, under different illumination conditions or when the object can be seen only partially. In this case, how can the object representations in memory be accessed and used?

Wood and Blair (2010) report findings where people use the mean value of the missing feature value obtained from previous experience. However, their experiments show that people do not always calculate the missing feature value. They should have an expectation that a feature is missing. Thus, they claim that it is not as simple as computing the mean value of previously encountered feature values but rather an inferential reasoning process about the identity of a feature.

Wood and Blair (2010) take into account correlations with other features, using known values to predict unknown ones and subsequently using the inferred values to generate an appropriate category response. Again, we have a situation where the visual system displays a complex pattern of behaviors that we believe only complex processes can produce. Such a pattern can be the result of a simple activation process. Wood and Blair point out that current computational models of categorization cannot explain this phenomenon. The reason for not being able to come up with such a model is that what we know about the structure of the neural populations has been very limited. However, some clues have started to appear, guiding us to better models of the visual system. What they offer as an explanation consists of generating a model to fit their findings. Again, without having a proposed system, they generate a model that only fits their data, which obviously would create problems when compared to an enormous set of proven effects in the literature.

2.4 Object representations in the contemporary models of object recognition and categorization

The question of how visual object representations are formed in memory has been addressed by various models from different disciplines including cognitive psychology and cognitive neuroscience. Even though the main objective of these models is not directly related to the representations of objects, they make specific claims about the structure of these representations. For example, models of object recognition aims to explain the mechanisms of how objects are recognized, but the formation of object representations is implicitly included in the specifications of the recognition mechanism.

In this section, models of visual perception and memory are grouped into three categories: convergence models, item-matching models and feature models. The basis for this categorization is the structure of the object representations adopted by the models. This is a different approach than the usual classification of these models which is based on the task accomplished by the model, like recognition, or categorization. Since the models are examined in terms of the representations they form, presenting them according to the representations provides a more structured analysis.

2.4.1 Convergence models

Visual object representations in the convergence models of object recognition and memory are distributed and have hierarchical representations which form during the training of the model at different levels of abstraction, usually from low-level stimulus features to conceptual representations. This transition from simple to complex representations is through convergence, which is a many-to-one mapping from low-level features, like bars and edges, to conjunctions of these simple features, like lines and shapes (Rolls & Deco, 2002).

A common property of these models is their emphasis on biological plausibility. The architecture of the models mimics their biological counterparts, and reference is made frequently to neuroscience studies. Examples of convergence models are VisNet (Rolls & Milward, 2000; Rolls & Stringer, 2006; Wallis & Rolls, 1997), RBC (Biederman, 1987; Hummel & Biederman, 1992), HMAX (Riesenhuber & Poggio, 1999), and PDP (Rogers & McClelland, 2005; Rumelhart & Zipser, 1985).

The processing in these models starts with low-level stimulus features, in contrast with psychological models which operate on more abstract entities. The architecture is almost always connectionist, where each incoming input to the network is used to update connection weights between input and output layers. The number of layers and nodes differ significantly in each model. Biological constraints, task constraints and stimuli constraints determine the layer structure. Low-level features in lower layers converge into more abstract representations at higher levels of the network. The top-most layer represents the objects whereas intermediate layers represent conjunctions of low-level features.

Even though the algorithm and the structure differ among the convergence models, they perform a similar computation: discovering regularities in the input space. Nodes in the output layers tune themselves to the co-occurrence of particular features in the input layers – nowadays often modulated by hidden layers that may considerably change the structure of the original input pattern.

There are a number of problems faced by convergence models. The first problem is invariant representations of objects. Biological visual systems can recognize objects from different viewpoints and under different lightning conditions. However, a hierarchical convergence model can only represent recurring conjunctions of features. Thus, models incorporate various strategies to deal with the invariance problem. A common approach is to train the network with images of the object from different viewpoints (Földiák, 1991). This way, the model discovers the properties of the object which are invariant to changing viewpoints.

VisNet

VisNet is a hierarchical, unified model of visual object recognition. Is has a strong neurological basis especially for the low-level processing of features. It has been implemented to test specific hypotheses about the formation of visual object recognition and the formation of memory. The model has two main components: a learning rule and a competitive output layer (Rolls & Deco, 2002; Rolls & Milward, 2000; Rolls & Stringer, 2006). The learning rule in this model is based on the main hypothesis that invariant object representations can be achieved by considering the trace decay property of the neurons. The trace decay property is the persistent firing of the neurons in a specific time frame depending on the presentation duration of the stimulus. For example, a stimulus presented for 16 ms results in a 100-400 ms persisting firing of neurons (Rolls & Deco, 2002). If the next stimulus is presented in this time frame, neurons will still be active. Thus, they will take role in the encoding of both the first and the second stimulus. Consider an object moving across the visual field. At each time point, the object will be at a different location. Since the neurons encoding the object at the first location will still be firing when the object is at the second location, they will also encode the object in this second location. Thus, these neurons will encode the object at every position. This will lead to a representation of the object independent of the specific location it has appeared in. This mechanism is embedded into VisNet through the trace learning rule.

In the trace learning rule, connection weights between the input and the output units are updated according to the decaying trace of previous activity on the output unit (Rolls & Deco, 2002):

$$\Delta w_j = \alpha \overline{y}^{\tau} . x_j$$

$$\overline{y}^{\tau} = (1 - \eta) y^{\tau} + \eta \overline{y}^{\tau - 1}$$

Where x_j is the j^{th} input to the neuron, y^t is the trace value of the output of the neuron at time step t, w_j is the synaptic weight between j^{th} input and the neuron, y is the output from the neuron, α is the learning rate between 0 and 1 and η is the trace value adjusted according to the presentation sequence length.

The trace learning rule is based on neurophysiologic properties but the implementation of the rule in VisNet has some biologically implausible additions, like normalization of the weights after the update. Rolls and Deco mention that they could also use a more plausible local weight bounding operation but he prefers normalization because of simplicity.

The competitive output layer is required to avoid redundant representations. According to Rolls and Deco, representing every view of the object is unnecessary and resource consuming. Thus, a single unit in VisNet represents multiple views of the same object/feature, and avoids other units to represent the same object/feature by inhibiting their activation.

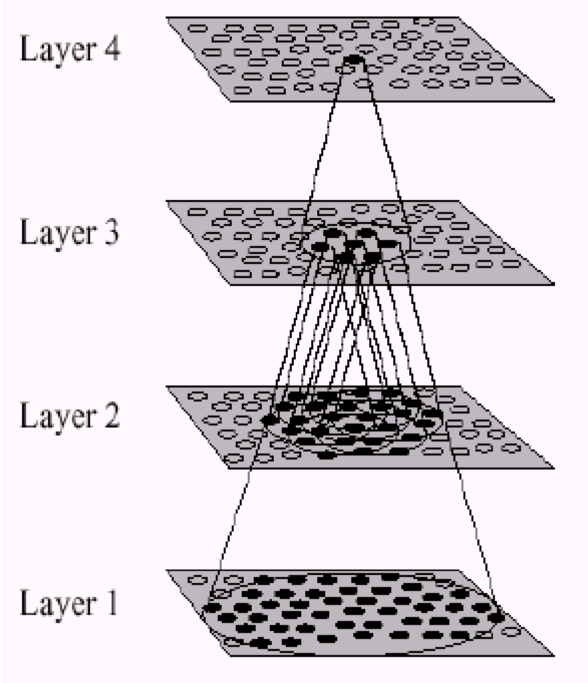


Figure 4 The depiction of the hierarchical structure of VisNet. The receptive field of the neurons increases from Layer 1 to Layer 4. Each output unit receives signals from a small subset of the input units, defined by a radius parameter (Rolls and Stringer, 2006, 44).

To enable this mechanism, each layer has horizontal inhibitory connections in addition to the vertical feed-forward activating connections. However, the radius of the inhibitory effect is restricted to a small value since the model should allow other units to represent different objects/features. The inhibition should only avoid multiple representations of the same object/feature.

Units in the competitive output layer are connected to units in the lower layer according to the regularities in the input. From Layer 1 to 4, units represent larger regions of the visual field (see Figure 4). Competition provides the ability to recognize partially occluded object. This way VisNet solves the missing feature problem. The occluded object activates existing representations to some amount. The representation with the highest activation wins the competition.

Output neurons receive input only from a small subset of the input neurons. This results in a hierarchical representation of the stimuli from a finer to a coarser representation of the stimulus. To enable this property, a definite value for the radius of the input range is defined.

A general evaluation of the VisNet

The success of the model results from the ability to demonstrate the possibility of formation of invariant object representations by a simple trace rule, starting from the very image in V1. Considering the trace learning rule, it is biologically plausible that motion helps formation of visual object representations. It makes use of continuous flow of information, rather than processing stationary images.

Findings from fMRI studies, however, indicate that viewpoint invariance is not always valid for object representations (Epstein et al., 2003). Thus, it is reasonable to claim that the different views are also encoded in memory. Most of the current models, including VisNet, prefer to encode different views into a single representation to create viewpoint invariant object representations. It seems more biologically plausible that both viewpoint-specific and viewpoint-invariant object representations exist in visual memory concurrently.

In addition, objects do not continuously move and rotate around. Infants are known to play with objects, rotate and observe them but this rotation is not necessarily long enough. The infant can take the object, look at other toys and look at the object from another viewpoint and so on. The visual system cannot rely on continuous input from the same object. It should also be able to make connections between object views from different time points. The system suggested by Rolls and Milward (2000) is a good candidate for explaining how discrete representations of features invariant to small scale changes in the visual field are formed. This way, the problem of how discrete feature representations are formed from a visual stream of input can be solved. However, it lacks the ability to integrate images from the object at different time points.

Rolls and Milward claim training VisNet with only a single subset of views is enough for building an invariant representation of the object. However, in real life, the visual system does not rely on only one experience with the object. Objects appear in different places at different points in time, sometimes for only a few seconds and sometimes much longer. VisNet does not have a mechanism for integrating new information about the objects and with only a trace rule it can only encode successive presentation of different views of the object.

2.4.2 Item-matching models

Item-matching models differ from convergence models in their level of explanation. They do not rely on neurophysiologic structures like convergence models do, but rather provide a mathematical explanation of the psychological phenomena. An object representation in memory is a point in the “psychological space” whose coordinates are the attributes of the object . Mathematical formulations are derived for explaining the relationship between these points in the psychological space using analytic geometry. The exemplar model of Nosofsky and Stanton (2005) and prototype theory of Rosch (1973) are two significant examples in this category.

Item-matching models differ among themselves by their theoretical claims about the computations taking place during classification of objects and the resulting mathematical formulations. For example, in the exemplar models, an object is classified to a category by computing the similarity of the object to all of the objects in the category. However, in the prototype models, only the similarity to the prototype object is computed. Even though the first one is more time-consuming it is able to explain various phenomena which cannot be explained by the prototype models. And of course there are phenomena which can be explained only by the prototype models. Thus, there are also hybrid models which take advantage of both model types.

Even though these are mostly models of categorization, they make strong assumptions/claims about representation of objects in memory. For example, the Generalized Context Model (GCM) is a model of categorization, but it is based on the claim that every object has a separate memory representation. Nosofsky, Little, Donkin, and Fific (2011) also mention this: “A central goal of exemplar models such as the GCM is not only to account for categorization but to explain relations between categorization and other fundamental cognitive processes, such as old–new recognition memory.” (p. 1). Since the feature frequency and similarity experiments in the present study are also old/new recognition experiments, there is strong correspondence between the suggested mechanisms of similarity computations and the findings of the experiments.

The Prototype Theory and Prototype Models of Categorization

Prototype theory was developed by Rosch (1975) to explain the formation of categories in culture, not in human memory:

> It should be noted that the issues in categorization with which we are primarily concerned have to do with explaining the categories found in a culture and coded by the language of that culture at a particular point in time. When we speak of the formation of categories, we mean their formation in the culture. This point is often misunderstood. The principles of categorization proposed are not as such intended to constitute a theory of the development of categories in children born into a culture nor to constitute a model of how categories are processed (how categorizations are made) in the minds of adult speakers of a language. ((Rosch, 1999) , p. 190).

Based on her experimental findings, Rosch concludes as follows:

- Membership is graded, i.e. some members are more typical of a category than others. Cutting and Schatz call this “analog” membership as opposed to a “digital membership”. An item can be member of more than one category, with varying degrees of membership. For example, a chair is a more typical member of the furniture category than a vase.
- Categories are centered on prototypes. A prototypical member contains “attributes” which are common to the items of the category and which are not common to items outside the category.
- There are basic level categories which display fundamental effects compared to subordinate and superordinate categories. Members of the basic level categories can be identified faster.

After the 1980s, attempts to build models of categorization in humans were made based on the prototype theory. However, Rosch (1978) was clearly against building models based on prototype theory, mentioning that the principles of the prototype theory should only be used as constraints to the process models, not for determining these models based on these principles. Lakoff (1987) supports Rosch’s point in that prototypes “do not constitute any particular theory of category learning”. (p 41).

Prototype Models

In a prototype model, a category is represented by a prototype which corresponds to the central tendency of the items in the category. A prototype can be a member of the category, or a more abstract construct like a feature bundle. It is usually calculated as the central tendency of the training items. There are various kinds of prototype models. As a simple example, a formula for computing the membership of an item to a category is calculated according to the formula, as in Casale and Ashby (2008):

$$P_{(A,B)}(A|x) = P(D_{xB} - D_{xA} > \varepsilon)$$

which states that the probability of an item to be classified into category A is the probability of the distance of the item to the prototype of the category B to be greater than the distance of the item to the prototype of category A. Here, $P_{(A,B)}(A|x)$ is the probability of item x to be classified into category A given two categories A and B. D_{xB} is the distance of item x to the prototype of B, and D_{xA} is the distance of item x to the prototype of A.

Generalized Context Model (GCM)

GCM is an exemplar based model of categorization (Nosofsky, 1992). It requires a training phase, where items are presented one by one to be encoded in memory. Every item has a separate representation. These representations are in the form of points in a psychological space. In this section the term "feature space" will be used instead of the "psychological space" since the dimensions of the psychological space corresponds to the features of the presented items. For example, an apple can be represented as a point in the feature space where the dimensions are color, size, shape, and taste.

Categories form naturally with each training item, since representations of similar items will be close to each other in the future space. Similarity between items is defined as the distance between their representations in the future space. Similar items will cluster into categories according to their feature values. However, not every dimension might be relevant to a category at every situation. Contextual cues might influence the relevance of the dimensions.

An attention mechanism is adopted in the GCM to incorporate contextual effects. Depending on the context, each dimension is assigned a weight. This weight is used in the similarity computation. Thus, the more contextually relevant the dimension, the more similar an item with the appropriate value for this dimension is. For example, if color red is important in the current context, its weight in the similarity computation is increased and the similarity of a red test item will increase.

2.4.3 Feature models

Object representations in feature models are collections of features connected to each other with different binding mechanisms. The level of explanation in these models is between the neurological level and the conceptual level. Some feature models are built to be compatible with findings from neuroimaging studies for their binding mechanism but some of them only model specific perceptual effects rather than the actual mechanisms of binding.

Some models represent objects hierarchically from low-level features to high-level features but the mechanism is not necessarily convergence. Others only define connections among features without imposing a hierarchy. But overall, feature models represent objects as collections of features whose connections result in various psychological phenomena. In the "object file" approach the connections between the features of an object are built when the object is first observed. The object file is like a file in which object features are stored.

There are two main mechanisms of binding adopted in these models. The first one is convergence, in which frequently co-occurring object features are stored in the same object file. The second mechanism is synchronization. The co-occurring features result in synchronous firing of feature neurons and this synchronized firing pattern is the representation of the object. Supporters of the synchronization mechanism claim that convergence is an inefficient process which requires multiple representations of the same object several times to detect the frequently co-occurring features. On the other hand, supporters of the convergence mechanism argue that synchronization is temporary and it cannot explain how long-term object representations are formed. Hommel tries to reconcile the two opposing views in this controversy and suggests a dual-processing mechanism where synchronous activation creates temporary bindings like a short-term memory representation, and convergence discovers regularities to represent invariant properties of objects, which is like a long-term memory representation of the object. Through this mechanism, he can explain phenomena that cannot be explained by a single-process mechanism.

The Theory of Event Coding (TEC)

In the TEC, a temporary object file is formed which encodes relationships between object features (Hommel & Colzato, 2009; Keizer, Colzato, & Hommel, 2008). When a similar object is presented, the corresponding object file becomes partially activated and updated with the new features. In other words, when a feature is presented, other features connected to it in the object file should be activated. However, if the feature is presented with features different from the features in the object file, there will be a conflict between the presented bindings and stored bindings which, as a consequence, would increase reaction time. "Higher order" bindings represent events, which binds objects, locations and actions together, hence the name "TEC".

Hommel and Colzato (2009) propose a dual-process model for feature binding (see Figure 5). The first process is "ad-hoc" binding which combines features arbitrarily. The second process is "conjunction detection" which discovers correlations among features and represents them. Ad-hoc binding is a synchronization process where features of the presented object create synchronous activation which temporally binds these features. On the other hand, the conjunction detection mechanism is based on previously learnt correlations stored in long-term memory.

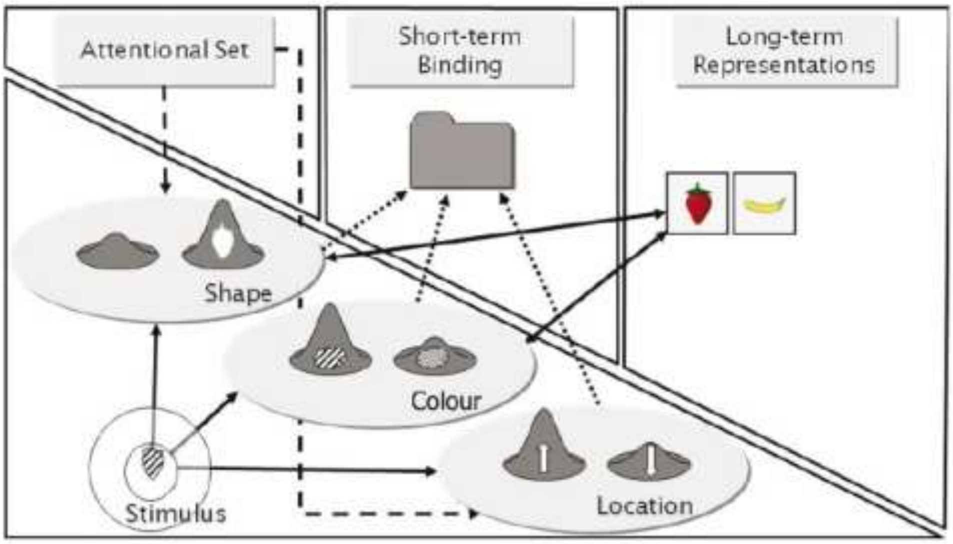


Figure 5 Effect of conjunction detectors on binding tasks. Conjunction detectors are depicted as “Long-term Representations” in the figure. (Hommel and Colzato, 2009, 126).

For example, presenting a red circle makes it easier to respond to a red circle presented afterwards compared to presenting a green circle. However, presenting a red triangle makes it harder to respond to a red circle compared to presenting a green triangle after a red circle. This means, totally unrelated features do not decrease performance as much as conflicting features. Conversely, related features increase the performance. Thus, even a single presentation of an object is enough to bind features of that object together, affecting the perception of the following object.

Hommel considers conjunction detectors as representation of objects in long-term memory. Building conjunction detectors requires learning whereas ad-hoc bindings are formed immediately without repeated exposure to the binding features. However, Colzato, Raffone and Hommel (2006) showed that real-world objects displayed higher binding effects. Hommel interprets this as an interaction between long-term memory of objects and temporary bindings.

2.5 Dissociations

The aim of dissociation experiments is to show that two processes differ by an experimentally manipulated variable. The effect of the variable is measured in both conditions. If an effect for one condition is found but no effect is found for the other condition, it is called a single dissociation. Mulligan and Lozito (2006) point to the ambiguity of single associations since these associations may not be reflecting a difference between two processes but only indicating a sensitivity difference between these processes. On the other hand, if an effect is found for one condition and an opposite effect for the other condition (e.g. an increase in the dependent variable in one condition and a decrease in the other condition) it is called a reversed dissociation.

Even though the dissociation approach has been heavily criticized, it is still one of the most common ways of examining various phenomena in cognitive psychology. It has been used to show the dissociations of STM and LTM, encoding and retrieval, familiarity and recollection, etc. Even though these dissociation studies present an enormous amount of findings, it is still not considered enough evidence for these dissociations since neuroimaging studies sometimes provide evidence against them. For example, reactivation studies show that the same regions are active during encoding and retrieval (Danker & Anderson, 2010). Again, this does not prove that the encoding and retrieval processes are essentially the same, but it suggests that at least the underlying mechanism might be common two both processes. Thus, one should consider the possibility that seemingly different processes can be two instances of the same mechanism, before arriving at a conclusion about the dissociation of these processes. The design of dissociation experiments should be devised accordingly. Producing more evidence for the dissociation of two processes does not eliminate the evidence from the neuroimaging studies. These experiments should also eliminate the possibility that the same mechanism could be responsible for the differential effects of the variables.

From the activation point of view, during encoding, some neural populations will encode the study items. When these items are presented during the test phase, the regions sensitive to the features of the stimulus will become active as well as other regions relevant to that stimulus to some degree. There would be a competition among these regions and the appropriate decision would be made according to the amount of activations in these regions. So, the same encoding regions will seem to be active during retrieval, and in addition, there will be some minor activations in less relevant regions, and some additional activation in the decision making regions.

With the above example, we would like to mention that the researchers should propose or support such neural mechanisms in order to discuss issues like dissociations of encoding/retrieval etc. Experimental manipulations should not be determined to show whether two things are different but to reason about the relations between these two things in the context of a larger system and then determine the factors to manipulate in the experiment and make predictions according to the proposed or supported mechanism. Even while talking about the 'basis' of these effects, they talk about response bias or familiarity change. A real basis for this effect would be the activations of neural populations causing the effect.

These memory processes should be considered as parts of a unified system. Without questioning the basic mechanisms, these discussions would produce only new evidence which would lead to yet other evidence without leading to an explanation of the actual mechanisms.

Of course there can be different levels of explanations, the neural level, the psychological level, etc. However, these should be in connection with each other to inspire and not to contradict each other. A psychological explanation contradicting neural phenomena would be unacceptable and a neural property not able to explain a psychological phenomenon would reveal a need for improvement.

2.6 Overlapping perceptual and action features in object representations

The link between perception and action is important for this dissertation since it is claimed that there are no separate object representations; instead, object features in different parts of the brain converge to layers of representations from more concrete to more abstract. Some of the most crucial convergence regions are the ones which receive connections from both visual object features and action features used in action planning. Daprati and Sirigu (2006) provide examples from brain lesion studies, reporting that identifying an object is not required for acting on the object. He also argues with the ‘visuomotor priming effect’ that implies the preparation of actions beforehand when the geometrical features of a presented object are perceived. This finding is frequently shown in affordance studies, which claim that object affordances are directly perceivable to the observer. Daprati and Sirigu mention connections from inferior parietal regions to temporal cortex as evidence for the integration between object properties and movement patterns. Likewise, auditory features may also have connections to the motor regions, for example, hearing the sound of a scissor cutting a piece of paper might activate motor planning of using the scissor. Shinn-Cunningham (2008) argues that auditory perception is based on almost the same principles as visual perception, in terms of phenomena like perceptual units of attention, top-down and bottom-up effects, and object individuation.

2.7 Summary and Contributions

A survey of literature on visual object representations has been presented in this chapter. State-of-the-art research clearly shows that objects are represented by a distributed set of features in different regions of the brain. These features can be visual or spatial properties and they are connected to each other so that when other cognitive processes need to reach the object representation, these connections activate the features belonging to that representation. Object representations are mostly illumination and perspective invariant, although some sensitivity to particular viewpoints is also observed. The representations differ in their intensity: the more intense they are, the better recognition they provide. Intensity of a representation is closely related with the amount of attention paid during formation of this representation.

Theories on object recognition and categorization also provide new hypotheses about object representations. Three groups of models incorporating different hypotheses about object representations were reviewed: convergence, item-matching and feature models. Even though structural differences among the representations involved were abundant, the main theme was the same: representations should enable recognizing both individual objects and regularities among perceived objects. Most models are successful at only one side; either representing individuals, or regularities. Other constraints on object representations include, but are not restricted to, enabling viewpoint and illumination invariance as well as identification of specific viewpoint and illumination dependent versions, graded category structure, and hierarchical structure from low-level features to complex object representations. The more constraints are obtained, the better models for the formation of visual object representations would be developed.

There are two main contributions of this study to the literature about formation of object representations. The first one is offering a new set of constraints that any model of visual object representations should satisfy. These constraints were obtained through two behavioral experiments explained in Chapter 3. The second one is a new model for the formation of visual object representations. This new model inherits many properties from the existing models, but has a broader scope in that both individuals and regularities can be represented using the same structures. This makes it possible to explain object recognition and categorization in a single framework. In addition, the model is built upon a framework which is based on findings from neuroscience studies, details of which are presented in Chapter 5. Employing an existing biologically-plausible framework instead of building a model with ad hoc structures made it possible to develop new predictions about the possible neural structures underlying object recognition.

CHAPTER 3

EXPERIMENTS

Two experiments were performed to investigate the effect of repetition frequency of features and feature combinations and similarity on the recognition of objects. In the first experiment the hypothesis that frequent repetition of particular features and feature combinations would increase “old” responses for objects which have the frequently repeated features (FRFs) was tested. In the second experiment, the hypothesis was that the similarity of objects to previously seen objects would not affect old/new responses when the repetition frequencies of features and feature combinations are equal. After the explanation of the methodology of the experiments, the results are presented and briefly discussed. A more detailed discussion of results can be found in Chapter 6.

3.1 Experiment 1 – The effect of feature frequency on old/new recognition of objects

Old/new recognition is a simple type of recognition where the participants are required to identify studied and unstudied objects. In the standard study/test old/new task, a set of items is presented to the participants during the study phase, and then another set of items is presented during the test phase where some items are from the study phase.

In the feature frequency experiment, the study/test old/new recognition paradigm was employed. Since the aim of this experiment was to test the effect of feature frequency on old/new responses, study and test items were constructed by manipulating the frequency of features and feature combinations. If the repetition frequency has a role in the formation of visual object representations, the visual object representations forming during the study phase should be affected by the manipulation of feature frequencies. Since old/new responses during the test phase are assumed to be based on object representations formed during the study phase, the feature frequencies should have an effect on the old/new responses. It was hypothesized that the increase in the frequencies of features and feature combinations increases the "old" responses.

3.1.1 Method

Participants. 20 adults participated in the experiment. The age of the participants ranged between 22 and 35 years. All participants were university graduates and had normal or corrected-to-normal vision. People who reported to be colorblind were not accepted to the experiment.

Stimuli. The study and test stimuli were created as follows.

Study stimuli

- There were four types of features: colour, shape, border and pattern. Each type had three values, as shown in Table 2. It was possible to create 81 objects using 4 features with 3 different values (3^4).
- Fifteen objects were selected among the pool of 81 possible objects, according to the following criteria: Solid black border and green color (pair 1) should repeat together on 5 objects (see Figure 6.a for an example of such an object). Diagonal line pattern and square shape (pair 2) should repeat together on 5 objects (see Figure 6.b). Other feature pairs should exist on 2 objects at most.
- Thus, FRFs were solid black border, green color, diagonal line pattern, and square shape, each repeating 7 times. Other features repeated only 4 times, e.g. 4 objects had blue color.
- These fifteen objects were created using the AutoShape tool of Microsoft Power Point. Objects had the same height (5 cm) and width (5 cm).

Test stimuli

18 objects were selected from the pool of 81 objects. 8 objects were from the study stimuli. The remaining 10 objects were unstudied objects which were not among the study stimuli. Objects had 0 FRF, 1 FRF or 2 FRFs. For example, an object which had 2 FRFs had green color and solid black border. An object which had 1 FRF had either green color or black border. An object without an FRF did not have these features at all. The number of objects under category is displayed in Table 3.

Setting. Computers at the Informatics Institute Computer Lab were used for the experiments. Stimuli were presented on a 19” widescreen LCD monitor by Microsoft Power Point software.

Table 2. Feature types and values used in the experiment

Colour	Red	Green	Blue
Shape	Square	Triangle	Circle
Border	Solid black	Dashed black	Coloured
Pattern	Dots	Diagonal lines	Shingle

Table 3. Number of objects for each level of the independent variables “study condition” and “number of FRFs” in the test phase.

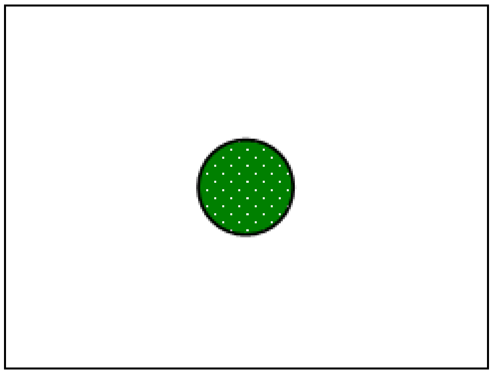
	Studied	Unstudied
Objects with two FRFs – pair 1	2	2
Objects with one FRF – pair 1	1	2
Objects with two FRFs – pair 2	2	2
Objects with one FRF – pair 2	1	2
Objects without any FRFs	2	2
Total	8	10

Experimental Design. There were two independent variables: study condition and number of FRFs. Study condition had two values: studied (1) and unstudied (0). Number of FRFs had three values: 0, 1 and 2. The dependent variable was the old/new score. It is the average of old/new responses given to the objects in a category. Categories are displayed in Table 2. This was a 2x3 repeated measures design.

Hypothesis.

H1. The number of FRFs on the test object should affect the average old/new score of the object.

(a)



(b)

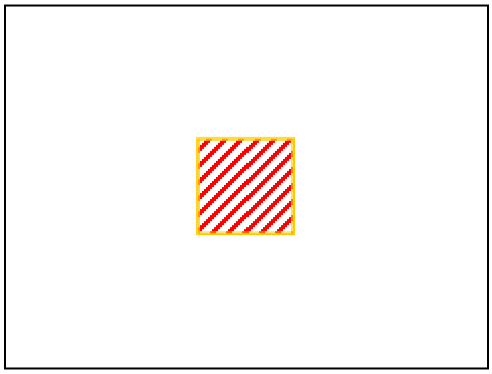


Figure 6. Example stimuli from the study phase of the first experiment. These objects include features that have high repetition frequency (a) Green color and solid black border (b) Oblique pattern and square shape

Procedure. The study/test old/new recognition paradigm was employed. In this paradigm, first, study items are presented to the participants. After the study phase, test items are presented. Participants are required to give an “old” or “new” response to each test object[1].

---

[1] Instead of the standard old/new responses, the participants were required to tell whether they “have seen” vs “have not seen” the item before. This is a more appropriate way of asking whether the

In this experiment, objects in the study stimuli were presented one by one using Power Point software. The center of gravity of the object was aligned with the center of the slide. Each Power Point slide was displayed for 2 seconds. Slide transitions were automatic. Two sets of 15 stimuli were presented

Immediately after the study phase, the test phase started. Each slide was displayed for 3 seconds. The order of the slides was reversed in half of the participants. As the subject responded to each slide, the experimenter noted +/- marks on a response sheet.

Before the experiment, participants signed an informed consent form. Experimental instructions can be found in APPENDIX F.

3.1.2 Results

Since there were two objects for each level of the independent variable, old/new scores were calculated by computing the mean response for each level for each participant. For example, if the participant responded with "old" to both objects, the old/new score was 1 ($response_1$= 1, $response_2$= 1, average($response_1$, $response_2$)= 1). If one of them was "old", and the other one was "new", the old/new score was 0.5 ($response_1$= 1, $response_2$= 0,

Table 4. Responses for the old/new recognition task. The numbers '0', '1' and '2' at the top of each column correspond to the number of FRFs on the object.

---

object is "old" or "new" in Turkish ("gördüm" is the word for "seen" and "görmedim" is the word for "not seen" in Turkish).

	Stimulus											
	Color and border repeated						Shape and pattern repeated					
	Studied			Unstudied			Studied			Unstudied		
Response	0	1	2	0	1	2	0	1	2	0	1	2
“Old”	35	28	37	9	16	27	35	36	32	9	17	28
“New”	5	12	3	31	24	13	5	4	8	31	23	12

average($response_1$, $response_2$)= 0.5). If both objects were “new” the old/new score was 0 ($response_1$= 0, $response_2$= 0, average($response_1$, $response_2$)= 0). Counts of old/new responses are displayed in Table 4.

Color and border. The effects of the two independent variables, study condition (studied, unstudied) and the number of FRFs (0, 1, or 2), were analyzed in a two-way repeated measures ANOVA. There was a main effect of study condition ($F(1,19)=46.77$, $p<0.001$, $\eta^2=0.7$), a main effect of the number of FRFs ($F(2,38)=13.57$, $p<0.001$, $\eta^2=0.4$) and an interaction between study condition * number of FRFs ($F(2,38)=3.57$, $p<0.05$, $\eta^2=0.2$). The mean old/new score was higher for the studied objects, and the main effect study condition implies that this was significant. In other words, participants could successfully remember the objects that had been presented to them before. The main effect of the number of FRFs shows that the old/new response of the participants was affected by the number of FRFs on the object. As the number of FRFs increased, the average old/new score increased. The third significant effect is the interaction effect. In Figure 7, the different patterns of responses for studied and unstudied objects can be seen. The number of FRFs did not affect mean old/new scores for the studied objects. However, for the unstudied objects, we see a totally

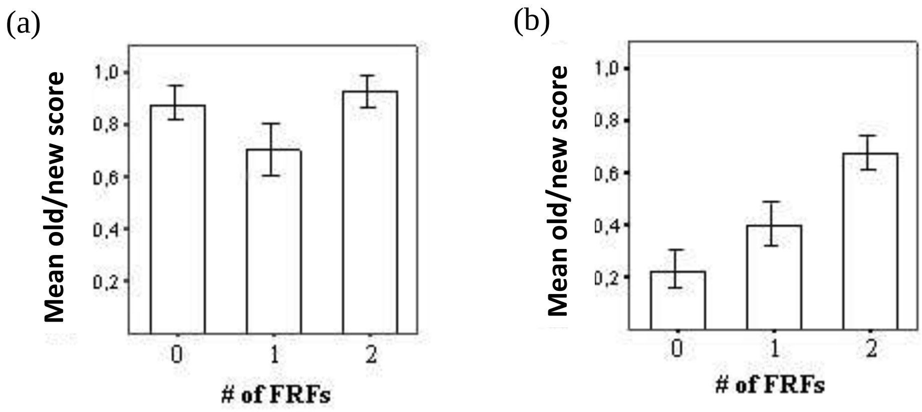


Figure 7. Mean old/new scores for the objects with zero, one or both of the features color green and solid black border. Error bars represent standard error. There was a main effect of study condition ($F(1,19)=46.77$, $p<0.001$, $\eta^2=0.7$[2]), a main effect of the number of FRFs ($F(2,38)=13.57$, $p<0.001$, $\eta^2=0.4$) and an interaction between study condition * number of FRFs ($F(2,38)=3.57$, $p<0.05$, $\eta^2=0.2$). (a) Studied objects (b) Unstudied objects

different picture. If the object had no FRFs, then most of the participants reported that they had not seen the object before. If the object shared only one of the FRFs, the mean old/new score doubled. Finally, if the object shared both of the FRFs, most of the participants reported that they had seen the object, although they had not.

Shape and pattern Likewise, for the second pair, the square shape and the diagonal lines pattern, the effects of study condition (studied, unstudied) and the number of FRFs (0, 1, 2) were analyzed with a two-way repeated measures ANOVA. There was a main effect of study condition ($F(1,19)=28.89$, $p< .001$, $\eta^2=0.6$), a main effect of the number of FRFs

[2] "η^2" denotes "partial eta squared".

($F(2,38)=5.67$, $p< .01$, $\eta^2=0.2$) and an interaction between study condition * number of FRFs ($F(2,38)=10.89$, $p< .001$, $\eta^2=0.4$). The mean old/new score was higher for the studied objects, and the main effect of study condition implies that this was significant. The main effect of the number of FRFs shows that the old/new response of the participants was affected by the number of FRFs on the object. As the number of features increased, the mean old/new score also increased. The third significant effect is the interaction effect. In Figure 8, the different patterns of responses for studied and unstudied objects can be seen. The number of FRFs did not affect mean old/new scores for the studied objects. For the unstudied objects, however, we see an effect of FRFs. If the object had no FRFs, then most of the participants reported that they had not seen the object before. If the object had only one of the FRFs, the average old/new score doubled. Finally, if the object had both of the relevant features, most of the participants reported that they had seen the object.

Effect of feature types on old/new responses. The aim of this analysis is to test whether there was a difference between effects of repeating the color/border pair and repeating the shape/pattern pair on the old/new judgment. The reason for doing this analysis is to be sure that the effect of FRFs on old/new decision is independent of the specific feature type, like color and border. Mean old/new scores for each pair are depicted in Figure 9. “p1” represents the feature pair green color/black border and “p2” represents the feature pair square shape and diagonal line pattern. For hits, we see a slightly different pattern for p1 and p2. For false alarms, old/new responses for p1 and p2 are almost identical. In this analysis we want to check whether the difference between p1 and p2 for the hits is significant. Two 2 (the number of FRFs: 1, 2) x 2 (feature pair: 1, 2) repeated-measures

(a) (b)

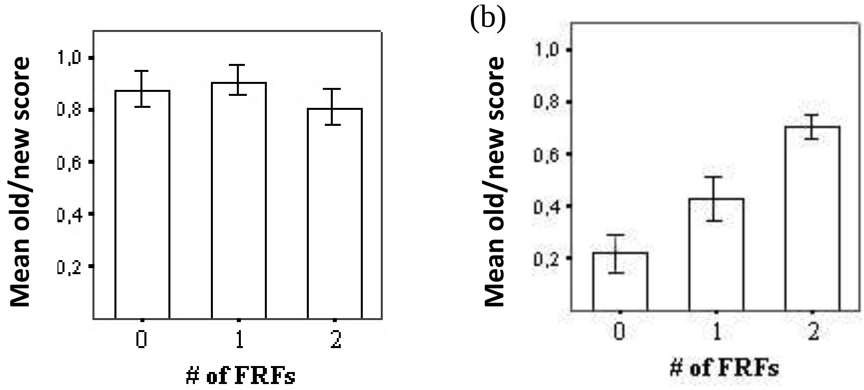


Figure 8. Average old/new scores for the objects with zero, one or both of the features square shape and diagonal lines pattern. Error bars represent standard error. There was a main effect of study condition ($F(1,19)=28.89$, $p< .001$, $\eta^2=0.6$), a main effect of the number of FRFs ($F(2,38)=5.67$, $p< .01$, $\eta^2=0.2$) and an interaction between study condition * number of FRFs ($F(2,38)=10.89$, $p< .001$, $\eta^2=0.4$). (a) Studied objects. (b) Unstudied objects.

ANOVA were performed separately for hits and false alarms. For false alarms, there was no significant difference.

For hits, there was an interaction effect between feature pair and the number of FRFs, $F(1,19)=4.65$, $p< .05$, $\eta^2=0.26$). The interaction effect showed that as the "old" responses increased with the # of FRFs for the objects with green color and solid black border, a decrease was observed for the objects with square shape and diagonal lines pattern.

3.1.3 Discussion

The first experiment aimed to test the hypothesis that the number of FRFs on the object affects the average old/new score for that object.

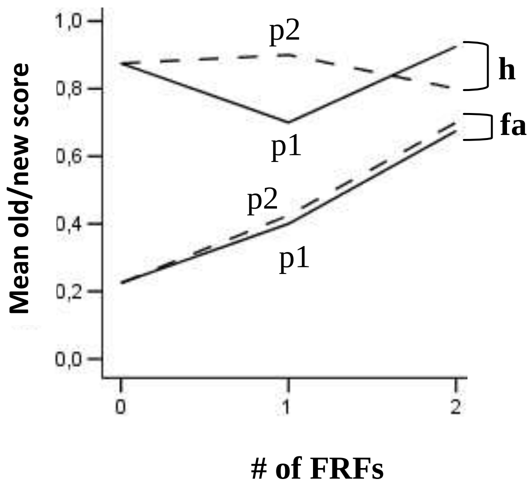


Figure 9. Old/new scores for hits and false alarms for two different feature pairs. h denotes hits and fa denotes false alarms. p1 represents the feature pair green color/black border and p2 represents the feature pair square shape and diagonal lines pattern. For hits, there was an interaction effect between feature pair and the number of FRFs, $F(1,19)=4.65$, $p< .05$, $\eta 2=0.26$).

This hypothesis was confirmed by the results: as the number of FRFs on the object increased, the average old/new score increased. This shows that the repetition frequency of object features during the study phase has an effect on the memory encoding of objects.

On the other hand, one can still claim that the difference between levels of the number of FRFs might be due to the difference between similarities of test items to the study items. Figure 10 shows mean old/new scores versus the similarity of test items to study items. There is an increase in the average old/new scores as the similarity of test items to study items increase. Thus, the second experiment was designed to check the possibility that similarities might have been responsible for the increase in the average old/new scores.

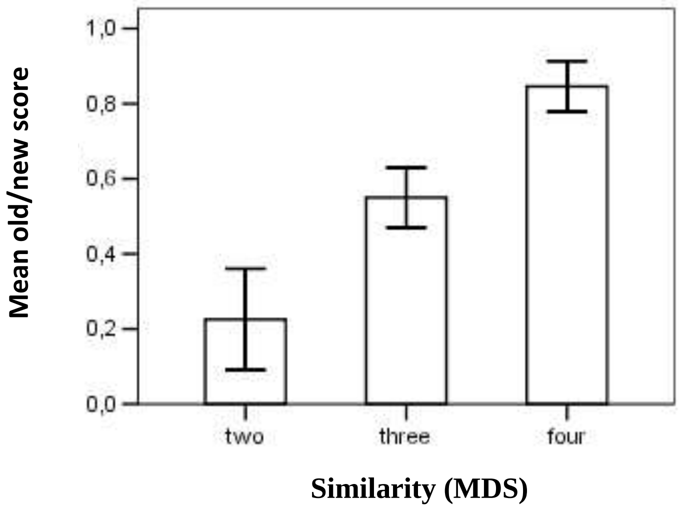


Figure 10. Mean old/new scores for test items which are two features similar, three features similar and four features similar (which are studied items) to the study items.

3.2 Experiment 2 - The Effect of Maximum Discrete Similarity (MDS) on old/new responses

The second experiment examined the effect of similarity on the old/new decision in a study/test old/new recognition task. We devised a mathematical formulation of the similarity measure adopted in this experiment. Similarity of each test object to the objects in the study phase was calculated according to this formulation. A repeated-measures ANOVA performed on the old/new responses with similarity measure as the independent variable showed that the similarity factor does not affect old/new responses in this setting. Reaction time recordings were analyzed according to the hypotheses derived from the Convergence-Divergence Zone (CDZ) framework (Meyer & Damasio, 2009). “Old” responses were significantly faster than “new” responses, supporting the hypothesis that giving a “new” response requires more evidence than giving an “old” response.

In our previous experiments, we had hypothesized that in a study/test old/new recognition task, high frequency of particular features during the study phase would increase "old" responses for objects which had these frequently repeated features (FRFs) during the test phase. The results indicated that unstudied objects with two FRFs received significantly more frequent "old" responses than unstudied objects with only one FRF. Studied objects were not affected by the FRFs. Besides, there was no difference between unstudied objects with two FRFs and studied objects. In other words, they appeared equally old to the participants. However, we were not able to attribute this effect to feature frequency only, since there was also the similarity factor which we had not controlled so far. Thus, in the current experiment, we aimed to test the effect of similarity only.

Object similarity. In our experiments, we employ the study/test old/new recognition task. In this task, study and test stimuli are presented in separate trials. In the first part, study objects are presented one by one. In the second part, test objects are presented. Test stimuli consist of studied objects and unstudied objects. The participant is required to give a binary old/new response for each test object. In this context, we devised two basic measures of similarity: discrete similarity (DS) and maximum discrete similarity (MDS). These similarity measures are explained below.

Discrete similarity. This measure is the number of common features between two objects. Given a set of features F and objects o_1 and o_2 defined over F, we can define discrete similarity as a function of o_1 and o_2 as follows:

Let $o_1 = (a_1, a_2, ..., a_n)$ and $o_2 = (b_1, b_2, ..., b_n)$ where $a_n, b_n \in F$

$$DS(o_1, o_2) = \sum_{x=1}^{n} f(a_x, b_x) \text{ where f(x,y) = } \begin{cases} 0 & \text{if} \quad x \neq y \\ 1 & \text{if} \quad x = y \end{cases}$$

$x = y$

f(x,y) is a function which takes two features as arguments and compares them. If two features are equal, then it returns 1. Otherwise it returns 0. The comparison continues until all features of the two objects are compared. The sum of the values of f(x,y) is the discrete similarity of the two objects.

Maximum discrete similarity. The DS measure of similarity defines the similarity of two objects. In order to manipulate similarity of a test item to study items in an old/new recognition task, an overall similarity measure is required. Given an object o_1 and a set of objects *O*, maximum discrete similarity is the highest number of common features between o_1 and each object in the set *O*. In order to calculate MDS, the discrete similarity between o_1 and each object in the set is computed. The highest value corresponds to MDS of the object o_1 to the set *O*. A mathematical formulation of the MDS measure is as follows:

$$MDS(o_1, O) = DS(o_1, \arg\max_{x \in O} DS(o_1, x))$$

Here, the argmax function returns the study item which carries $DS(o_1, x)$ to its maximum value. The maximum value of $DS(o_1, x)$ means the highest possible similarity of test item o_1 to an item in the study stimuli. Thus, the item returned by argmax is the most similar study item to the test item o_1. As a result, $MDS(o_1, O)$ is the similarity of object o_1 to the most similar item in the study stimuli. This measure is important since the most similar study item is like a match to the presented test item. This measure will be used to manipulate similarity of test items to study items.

There could be other measures of such similarity, like the sum of DSs (SDS), instead of the maximum similarity. We preferred MDS as the similarity measure and kept SDS constant. This preference results from theoretical considerations. The most similar test object to the study objects is a test object identical to an object in the study phase. If similarity was about the total number of common features with all the study objects, a test object which has common features with various objects in the study phase, without being identical to any of them, could be more similar to the study stimuli than an object which is identical to a single study object but has no common features with another object.

Moreover, one can also advocate continuous similarity instead of discrete similarity. Discrete similarity is computed by comparing objects feature by feature and checking if the compared features are equal. This may not be useful for comparing real-world objects. For example, in our experiments, there are only three values of the color feature: red, green and blue. In reality, the color values are graded. Two objects can be both red, but their intensities can differ. In that case, a binary comparison may not be suitable. The fuzzy set theory can be very suitable for this type of comparison. However, we do not include such a definition in this thesis, since it is not directly related to our hypotheses about visual object memory.

3.2.1 Similarity Experiment

In the previous experiment, the effect of feature frequency on old/new responses was tested. However, there was a similarity factor confounded with the feature frequency. In the feature frequency experiment, for objects with 3 similar features, only those with 2 FRFs had similar percentages of "old" responses as studied objects – whereas the percentage of "old" responses for those with only 1 FRF was as small as the percentage of "old" responses for objects with only 2 similar features. In the current

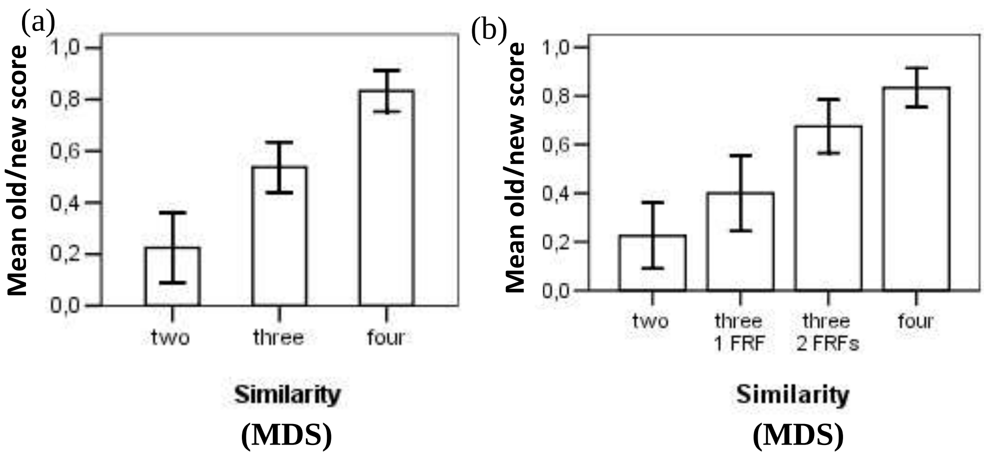


Figure 11 Mean old/new scores for levels of similarity (MDS) in our previous feature frequency experiment. Error bars represent standard error of the mean. (a) Number of FRFs is hidden, only MDS values are shown. (b) Here, for the same data, the number of FRFs is shown below the MDS values for three-similar objects. As can be seen from the graph, the increase in three-similars is actually a result of feature frequency. Objects with 1 FRF are not significantly different from two-similars. And objects with 2 FRFs are not different from four-similars.

experiment, we isolated the similarity factor by keeping feature frequencies of features constant to see whether we would obtain the same result for the effect of similarity.

We employed the same study/test old/new recognition task as in the previous experiments. The study and test stimuli were reconstructed according to the similarity factor. We used MDS as the similarity measure. The details of how the objects were selected were explained in the Method section. This time, the frequency of features was kept constant among the study stimuli. The participants gave binary old/new responses for test objects. E-prime software was used to randomize stimulus presentation and to record reaction times (RT).

3.2.2 Method

Participants. 22 adults from a Community Health Center in Ankara participated in the experiment. They were aged 33.6 years on average. All of them were right-handed.

Setting. Stimuli were presented with E-Prime software on a 15-inch laptop monitor. Responses were collected through the keyboard. The response to the recognition task was given with the right hand using the index and middle fingers. The experiment was run on Microsoft Windows XP.

Stimuli. Four feature types were used to create objects (see Table 5).

Table 5 Feature types and values

Color	**Shape**	**Pattern**	**Border**
Red	Square	Dotted	Solid
Green	Triangle	Spiral	Dotted
Blue	Circle	Line	Dashed

The procedure for creating study and test stimuli:

- All possible combinations of four feature types were created in Microsoft Power Point (3x3x3x3=81 objects).
- Study stimuli were selected among these 81 objects according to the following constraints:

- Each feature value should appear exactly the same number of times e.g. color green should be as frequent as spiral pattern.
- Three feature values should not repeat together, e.g. there should be only one object with green color, square shape and spiral pattern.

Only 9 objects satisfied the above constraints. Each feature value appeared three times, e.g. regarding color feature, there were three red objects, three blue objects and three green objects in the study stimuli.

- Test stimuli were selected among the pool of 81 objects as follows:
  - The Maximum Discrete Similarity (MDS) of every object in the pool to the set of study objects was calculated. In mathematical terms, this corresponds to calculating MDS(object, StudyStimuli) for each object in the pool where StudyStimuli is the set of study objects.
  - Objects were classified into three groups according to their MDS values: MDS_2 (18 objects), MDS_3 (54 objects) and MDS_4 (9 objects).
  - Four objects from MDS_2, four objects from MDS_3 and eight objects from MDS_4 were selected randomly by E-prime software during the experiment. This led to an equal number of studied and unstudied objects (objects in MDS_2 and MDS_3 are unstudied objects whereas objects in MDS_4 are studied objects).

Hypothesis.

H2. Similarity of a test item to study items would not affect the average old/new score for that item.

Procedure. First, two cycles of 9 study objects were presented in random order. Each object was presented in the middle of the screen for 3000 ms, followed by a 1000 ms fixation screen. Then, test objects were presented. Participants pressed the ‘space’ key to advance to the next object. They gave their response by pressing the ‘n’ key for ‘new’ and the ‘m’ key for ‘old’.

Participants were instructed verbally at the beginning of the experiment. Before the test phase, they were told to respond as quickly and as accurately as possible. It was emphasized that both speed and correctness were equally important.

Design. The experiment employed a within-subjects design with one independent variable, MDS (two, three, four). The MDS variable inherently represented the study condition variable, where objects with MDS=2 and MDS=3 were unstudied objects and objects with MDS=4 were studied objects. The test phase consisted of 16 objects, with 8 studied and 8 unstudied objects, presented in randomized order. The dependent variables were mean RT and mean FR for each level of MDS.

3.2.3 Results

Participants gave old/new responses for test objects by pressing the ‘N’ key for ‘new’ and the ‘M’ key for ‘old’ objects. In the data file these responses were re-coded as 0 for “new” and 1 for “old” response. For each participant, the mean old/new scores for each level of MDS was calculated.

Old/new scores. The effect of MDS on old/new scores was tested with a repeated-measures ANOVA. The independent variable MDS had three levels: two, three and four . The dependent variable was the mean old/new score for each participant for each level of MDS. Mean old/new scores varied between 0 and 1. There was no significant difference between levels

Table 6 Means and standard deviations of old/new scores (ONS) for each level of MDS.

	Maximum Discrete Similarity (MDS)					
	two		three		four	
	M	SD	M	SD	M	SD
ONS	0.75	0.26	0.74	0.25	0.85	0.16

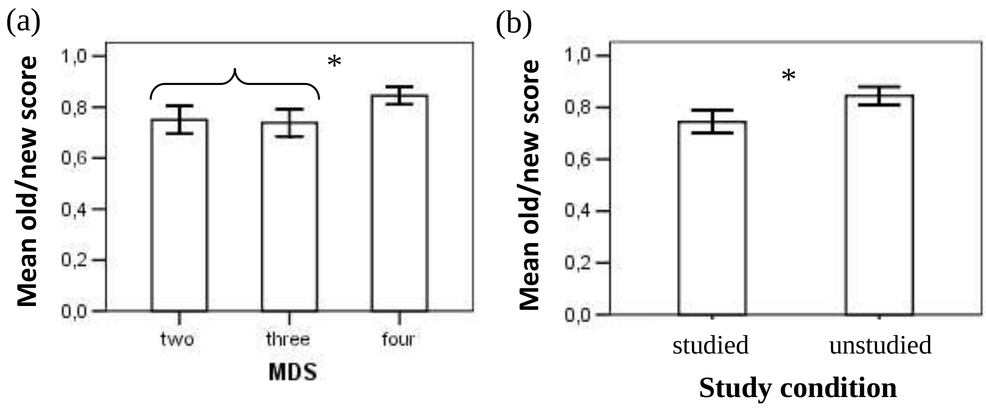


Figure 12 Mean old/new scores. (a) Mean old/new score for each level of MDS. No significant difference between groups. (b) Mean old/new score for studied and unstudied objects. “New” is an average over MDS “two” and “three”. Studied objects had significantly higher “old” response percentage than unstudied objects, $t(21)= 2.71$, $p<0.05$, $r=0.26$. Error bars represent the standard error of the mean.

of MDS in terms of old/new scores, $F(2,42)=2.37$, $p>0.05$ (see Figure 2-a). On the other hand, studied objects received significantly higher “old” response percentages than unstudied objects, as revealed by a Helmert contrast between MDS=4 and the other two levels (MDS=2 and MDS=3), $F=7.36$, $p<0.05$, $\eta^2=0.26$ (see Figure 12).

3.2.4 Discussion

In this experiment, the isolated effect of similarity (in terms of MDS) on old/new responses was tested. The frequency of features was kept constant among the study stimuli. MDS, only the similarity of test objects to study objects, was manipulated. Participants gave binary old/new responses for test objects.

As in our previous feature frequency experiment, similarity of test objects to the study objects (MDS) did not affect old/new responses. Actually, participants found objects very similar to each other and they could hardly distinguish one object from the other. This was expected since objects were combinations of four features and objects differed from each other by two features at most. Thus, they found most of the objects ‘old’. However, they were able to differentiate studied and unstudied objects, as revealed by the significant difference between average old/new responses for studied and unstudied objects (Helmert contrast). This is important because it indicates that representations of studied objects could form in memory during the study phase.

3.3 Summary of Experimental Results

Evidence from the similarity and feature frequency experiments showed that feature frequency not only increases the impact of frequently repeated features on old/new responses, but also suppresses the impact of infrequently repeated features. A general tendency towards giving an ‘old’ response was observed in the similarity experiment, causing homogeneity of responses for all levels of similarity. Feature frequency distorts this homogeneity by decreasing ‘old’ responses for unstudied objects which have infrequently repeated features. “Old” responses for studied objects remain unaffected, though, indicating that the memory for studied objects is preserved.

In the next chapter, the relation of the present findings to existing theories/models of visual object memory and categorization will be examined. This examination also aims to see to what degree existing models satisfy the constraints imposed by the experimental results of the present study. After the proposed model of the present study is presented in Chapter 5, these results will also be used to validate the model, and as a basis for comparison with selected models in the literature.

CHAPTER 4

IMPLICATIONS OF THE EXPERIMENTAL RESULTS FOR EXISTING THEORIES OF VISUAL PERCEPTION AND MEMORY

The question of how visual object representations are formed in memory has been addressed by various models from different disciplines including cognitive psychology and cognitive neuroscience. Even though the main objective of these models is not directly related to the representations of objects, they make specific claims about the structure of these representations. For example, models of object recognition aims to explain the mechanisms of how objects are recognized, but the formation of object representations is implicitly included in the specifications of the recognition mechanism.

In this chapter, models of visual perception and memory are evaluated with respect to the experimental results presented in the previous chapter. The models are grouped into three categories: convergence models, item-matching models and feature models. The basis for this categorization was the structure of the object representations adopted by the models. This is a different approach than the usual classification of these models which is based on the task accomplished by the model, like recognition, or categorization. Since the models are examined in terms of their explanations for how representations form, presenting them according to the representations provides a more structured analysis.

4.1 Findings from the feature frequency and similarity experiments with respect to the visual memory literature

The experiments presented in the previous chapter provided the following findings:

- In the similarity experiment, participants were shown a number of study items in which the repetition frequencies of features were equal. The similarity of test items to study items was controlled. When the test items were shown, participants responded with an "old" or "new" response to each item. The percentages of responses are shown in Table 7. The percentage of "old" responses to studied items was significantly higher than the percentage of "old" responses to unstudied items ($F=7.36$, $p<0.05$, $\eta^2=0.26$). There was no significant difference between levels of similarity.
- In the feature frequency experiment, participants were shown a number of study items in which the repetition frequencies of features were manipulated. When the test items were shown, participants responded with an "old" or "new" response to each item. The percentages of responses are shown in Table 8. The percentage of "old" response given to studied items was significantly higher than the percentage of "old" response given to unstudied items

Table 7. Percentage of old and new responses in the similarity experiment for each level of similarity, i.e. DS_4, DS_3 and DS_2.
H: Hits, M: Misses, FA: False Alarms, CR: Correct Rejections

Response \ Actual	studied	unstudied	
	DS_4	DS_2	DS_3
"old"	85% (H)	75% (FA)	74% (FA)
"new"	15% (M)	25% (CR)	26% (CR)

Table 8. Percentage of old and new responses in the feature frequency experiment for each level of FRF, i.e. 0, 1, 2. Level 0 denotes items without FRF, level 1 denotes items with one FRF, and level 2 denotes items with 2 FRFs.

Response \ Actual	studied			unstudied		
	0	**1**	**2**	**0**	**1**	**2**
"old"	88%	80%	86%	23%	41%	69%
"new"	12%	20%	14%	77%	59%	31%

(F(1,19)=46.77, p<0.001, η^2=0.7). In addition, there was a main effect of the number of FRFs (F(2,38)=13.57, p<0.001, η^2=0.4) and an interaction between study condition * number of FRFs (F(2,38)=3.57, p<0.05, η^2=0.2).

According to the hypotheses of this study, the more frequent repetition of some features as compared to other features during the study phase would increase the “old” responses for items having these frequently repeated features during the test phase. This hypothesis was supported by the results of the repetition experiment, as shown in Table 8. There was an unexpected effect, though: the interaction effect between study condition and the number of FRFs on the item, which indicated that, the effect of FRFs differed for studied and unstudied items. In fact, responses for studied items were not affected by FRFs: most of the participants gave “old” response for studied items. On the other hand, “old” responses increased for unstudied items as the number of FRFs on the test item increased. It seemed that FRFs made unstudied items look like studied items.

Even though the results of the repetition experiment were highly significant, there was still a possibility that similarity of test items might affect old/new responses. The similarity experiment was designed to test this possibility. The results were surprising: the number of “old” responses was high in general, even for unstudied items. The results of this experiment showed that the number of FRFs was not raising “old” responses for unstudied objects, but actually <u>decreasing</u> “old” responses for items <u>without</u> FRFs.

These findings indicate that humans are not very good at differentiating items with similar features. When they see an item similar to a previously presented item, most of them report that they had seen the item before. The measure of similarity in this study was the number of common features between two items. Similarity of test items to the study items is a factor which affects retrieval of the previous memories. It is not related to the encoding of object representations, since it is manipulated on test items.

The literature about the effect of similarity on retrieval of visual memory items is quite poor. Most studies concentrate on verbal items, and examine structural and surface similarities. Models of similarity for visual items are usually constructed in the scope of computer science, to build artificial systems. Existing studies on visual objects in the domain of cognitive psychology concentrate on changing pose, rather than structural properties of the objects (Edelman, Cutzu, & DuvdevaniBar, 1996). There is one specific study on similarity of objects by Shepard (1957): he found that the generalization among a set of objects increases as the similarity of objects increase. He defined similarity as the distance between objects in a representational space. In his experiments, generalization decreased exponentially as the distance between objects in this space increased.

Some models of categorization, also with verbal stimuli, treat similarity of items as their distance in a defined space. The name of the space varies among studies, such as psychological space, etc. These models differ mostly in terms of how they use this similarity measure in recognition of objects and categories. One of the major studies which use similarity information to determine the category of an item defined by multi-dimensional feature values is the Generalized Context Model (GCM) by Nosofsky and Johansen (2000). In this model all exemplars are encoded separately. As a new item is introduced to the model, for each category, the distances between the new item and all members of the category are computed and summed. The new item is assigned to the category with the smallest sum of distances. This model is also capable of explaining familiarity decision: the distance between the new item and all the stored items are computed and the closest item is retrieved as the match. Since this model has direct claims about similarity computations and exemplar-based item representations, it will be discussed in a separate section in this chapter with respect to the experimental results obtained in this study.

Another category of models with similarity computations in a representational space is the category of prototype models. In these models, similarity is computed only between the new item and representative items of the categories. A representative item of the category is called a prototype, an item which has features appearing on all category members but not on members of other categories. Since these models have been dominant in the categorization and recognition literature, they will also be examined separately in the following sections, with respect to the data from behavioral experiments.

In addition to the similarity of test items to the study items, the similarity of study items among themselves can be responsible for the false recognition of unstudied test items. The low discriminability of study items from each other might affect encoding of these items so that recollection of individual features becomes difficult. However, the significant difference between "old" responses for the studied and unstudied items does not support this hypothesis. If low discriminability had influenced the encoding of individual features, responses for studied items should have also been affected. The effect was observed only for unstudied items, with increasing "old" responses as the number of FRFs on the test object increased.

Examining the effects of feature repetition frequency is not common among the studies with visual stimuli. Featural representations are usually examined using artificial neural network models, which are computational, rather than behavioral studies. Obtaining findings from behavioral experiments is hard, since the manipulation of features requires automated presentation of the stimuli, controlling the number of presentation for each feature. Besides, all the items should be constructed either manually from a diverse set of features, or automatically, by writing a computer program to produce items from the specified features. Thus, there are very few studies which examine feature repetition effects using behavioral experiments.

4.2 Correspondence between findings of the present study and the false memory literature

The studies most relevant to the effects observed in the similarity and feature frequency experiments appear in the false memory literature. The finding that feature frequencies affect object recognition rather than similarity is compatible with Smith et al.’s (2000) finding that false memories are caused by output dominance rather than typicality. Note that in the similarity experiment, there was no significant difference between old/new responses given to different levels of similarity, where repetition frequencies of features were equal during the study phase.

An interesting parallel between the first experiment of Smith et al. (2000) and the present feature frequency experiment is that the effect of similarity (a measure of typicality) and feature frequency (increases accessibility) on responses are intermixed, even though they used verbal stimuli (presented on the computer screen) rather than visual stimuli. Smith et al. state that typicality and accessibility (in terms of output dominance) were found to be correlated so they performed a second experiment and found that the effect was due to accessibility, not typicality. In this study, the similarity experiment was performed with the same motivation. When feature frequencies were kept equal, there was no difference between levels of similarity.

The discussions of the effects of typicality on old/new responses are based on the prototype theory of Rosch (1973). According to the prototype theory, memberships of category members are graded, and people can state the most typical item of a category. Even though the prototype theory does not have a direct relation to the experiments of the present study since the task in the experiments were recognition rather than categorization, the findings in the false memory literature indicate that category effects can be responsible for the observed false memories. Considering the parallels between the findings

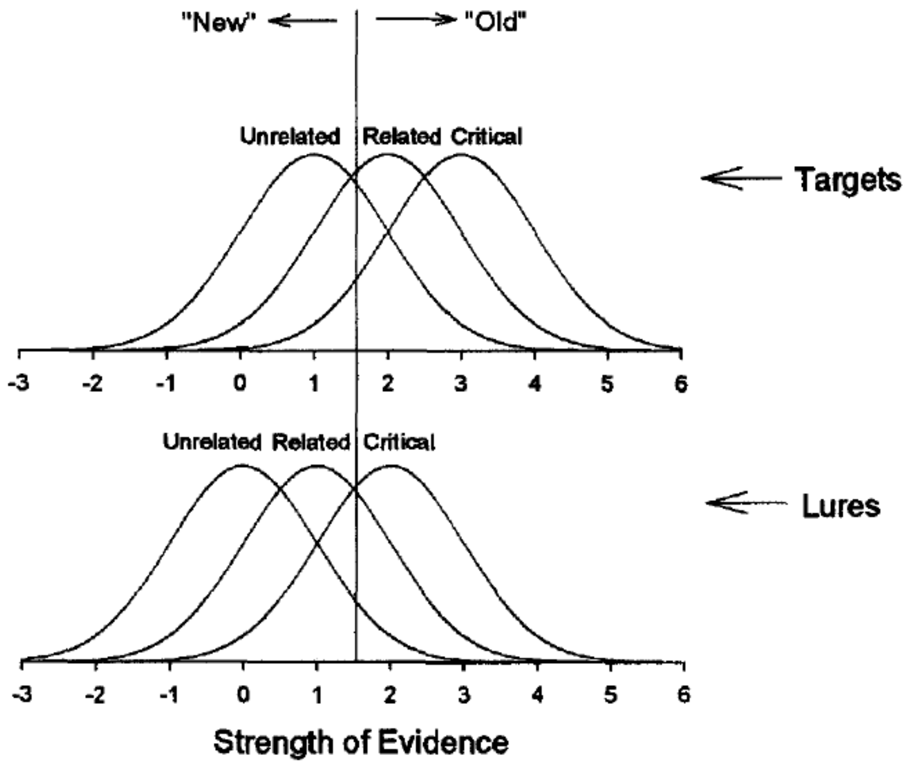


Figure 13. Predictions of the false memory model of Roediger and McDermott (1995), according to Wixted and Stretch (2000), p.370.

of Smith et al. (2000) and findings of the present study, it is possible that the patterns observed in participants' responses could be a result of the formation of category representations during the study phase because of the regularities in the study stimuli. These categories could have affected old/new responses during the test phase, according to the graded category membership of the test items. Whether typicality or output dominance is a better measure of graded category structure is another issue. According to Smith et al.'s findings, output dominance is more important. The experiments of the present study supports Smith's findings, if similarity is considered as a measure of typicality and feature frequency is seen as a factor affecting accessibility (and output dominance in turn).

Wixted and Stretch (2000) compare three models of false memory. The first one is a mathematical model of Roediger and McDermott's findings.

Table 9. Average values for item types, as predicted by the Roediger and McDermott (1995) model, computed by Wixted and Stretch (2000), p.370..

Critical target (CT):	$S_{CT} = 1 + 2 = 3$
Related target (RT):	$S_{RT} = 1 + 1 = 2$
Unrelated target (UT):	$S_{UT} = 1 + 0 = 1$
Critical lure (CL):	$S_{CL} = 0 + 2 = 2$
Related lure (RL):	$S_{RL} = 0 + 1 = 1$
Unrelated lure (UL):	$S_{UL} = 0 + 0 = 0.$

The second one is a modified version of the first model, by Miller and Wolford. The third one is a decision-based model, which is supported by Miller and Wolford. In the first model, the memory strength of a test item is computed as follows:

Strength of item (S) = Presentation (P)+ Other items (A)

The strength of a test item depends on the presentation of the item itself plus the associative effect caused by the presentation of other items. This model has interesting predictions. For example, if an item is not presented during the study phase, but some associated items are presented, there will be no contribution from P, but there will be contribution from other items, thus increasing A. On the other hand, the strength of an item is determined by its own strength only, if an associated item is not presented. Thus, the strongest item will be the one that is presented during the study phase, together with several associated items. The predictions of this model for the DRM effect are displayed in Table 9 and Figure 13.

When these values are computed for the results of the similarity and feature frequency experiments, the following percentages are obtained:

Studied objects with 2 FRFs: 86 %
Sudied objects with 1 FRF: 80 %
Studied objects with 0 FRF: 88 %

Unstudied objects with 2 FRFs: 69 %
Unstudied objects with 1 FRF: 41 %
Unstudied objects with 0 FRF: 23 %

If these percentages are mapped onto the 0-3 range, the following values are obtained:

Studied objects with 2 FRFs: 3
Studied objects with 1 FRF: 3
Studied objects with 0 FRF: 3

Unstudied objects with 2 FRFs: 2
Unstudied objects with 1 FRF: 1
Unstudied objects with 0 FRF: 0

The predictions of the Wixted-Stretch interpretation of the Roediger-McDermott (1995) model are perfectly matching the results of the feature frequency experiment for studied objects with 2 FRFs and all unstudied objects. However, they differ for studied objects with 1 and 0 FRF. No matter what the relation of the item to the study items is, it was rated as "old" if it was studied.

There are conflicting results in the literature about the false memory effects on studied items. For example, there is a phenomenon called "mirror effect" which is the interaction between the item being studied or unstudied and the frequency of the item. Increase in item frequency decreases "old" responses for studied items (in other words, decreases hits) but increases "old" responses for unstudied items (increases false alarms) (Glanzer, Adams, Iverson, & Kim, 1993). In the feature frequency experiment, the same interaction effect was found, but there was no decrease in “old” responses for studied objects (no changes in hits).

Smith and Hunt (1998) argue for some modality effects on false memory. They construct stimuli by producing visual versions of the original auditory stimuli from the experiment of Roediger and McDermott (1995), and compare responses to these visual stimuli with a control group to whom the original auditory stimuli were presented. They were unable to obtain false memory effects for visual items. Their result contradicts the findings of the similarity and repetition experiments where false memory effects were obtained for visual stimuli. Actually, they cannot provide an account of their results, and speculate that visual memories could be much stronger than auditory memory resulting in better memory performance, with lower false memory effects. However, as the similarity and feature frequency experiments demonstrated, false memory effects do occur also for the visual stimuli, and they show similar patterns as in other studies in the literature. The reason for the difference between this study and their study might be the method of stimuli construction. In both the similarity and feature frequency experiment, items were constructed by calculating feature frequencies and similarities of study and test items. Their stimuli were visual versions of the originally auditory stimuli. The selection of images might have been inappropriate as a depiction of the original stimuli.

Roediger and McDermott (1995) report that participants were very confident in their responses for falsely remembered unstudied items. The same pattern was observed in the similarity and feature frequency experiments: the more FRFs the object had, the more the participants were confident about their responses. Roediger and McDermott interpret this as an evidence for conscious recollection rather than a sense of familiarity.

Finally, the model proposed by Underwood (1965) and other more recent theories based on this model states that false memory is a result of associations activated during encoding. However, this cannot explain the findings of the similarity and feature frequency experiments. Frequencies of features during the study phase affected responses during the test phase. While Underwood uses names of everyday items, the stimuli in the similarity and feature frequency experiments were artificial and they were not represented in memory before the experiment.

4.3 Correspondence between findings of the present study and models of object recognition and categorization

In this section, selected models of object recognition and categorization are examined in terms of the results of the behavioral experiments of the present study.

4.3.1 Convergence models

The model which will be discussed in terms of the results of the feature frequency experiment is VisNet by Rolls and Milward (see section 2.4.1). It is a typical example of convergence models with its hierarchical layer structure inspired by the structure of the primary visual cortex and inferotemporal cortex. It claims to explain various phenomena regarding object recognition and visual memory. It is well-recognized in the object recognition literature and one of the most comprehensive models which is biologically plausible from many aspects.

Feature Frequency Effects in VisNet

In our experiments, we presented stationary objects one at a time. VisNet is based on the assumption that invariant object representation can be learned by receiving continuous input from the object during a specific transformation. Since the objects in the feature frequency experiment appear only once, the main capability of the VisNet, which is to represent invariant object representations, is disabled. However, Rolls and Milward (2000) claim that VisNet is a biologically plausible architecture which can be the basis of many perceptual and memory processes. Thus, an analysis of the VisNet when the stimuli from the feature frequency experiment are presented can provide insight into the behavior of such hierarchical models. Comparing these behaviors with the actual data from the experiment can provide evidence for the psychological validity of these models.

Stimuli from the feature frequency experiment as input to VisNet

Learning in VisNet corresponds to updating connection weights according to the trace learning rule. When the first object in the study phase of the feature frequency experiment is presented, filters in Layer 1 compute low-level stimulus features like bars and edges. Firing rates of Layer 2 neurons which receive input from Layer 1 increase, and connections of these neurons with Layer 1 neurons which send input to them are strengthened. Similarly, firing rates of Layer 3 neurons which receive input from Layer 2 increase, and connections of these neurons with Layer 2 neurons which send input to them are strengthened. After Layer 4 neurons update their connection weights with Layer 3, the hierarchical representation of the first object is formed. With this mechanism, structurally similar simple features combine to form more complex features and then to form representations of objects.

When the second object is presented, VisNet would show a tendency to connect the second object to the representation of the first object, since the trace from the first object will still remain active. This means, if the second object is similar enough to fire the neurons representing the first object, they will be encoded in the same representation in VisNet. This is a serious weakness of VisNet: whether the new stimulus is another viewpoint of the same object or totally new object should be explicitly told to VisNet. Rolls and Milward suggest that when the new object is not strongly correlated with the first one, the trace rule can be reset so that the new object would not be encoded by the previous representation and a new representation is formed. This would also lead to new problems, like how much correlation is needed to determine a new object, so Rolls and Milward leave this issue as a future work. Thus, even though VisNet is claimed to be a general-purpose object recognition system, it is only a mechanism of how invariant representation of objects can be learned from successive input from the object during its transformation.

VisNet as a multiple-layer self-organizing map

When the trace rule is omitted, VisNet becomes a self-organizing map with multiple layers each self-organizing itself according to the regularities in their input. This structure is more relevant to the data in the feature frequency experiment. Given the study stimuli, the network discovers the regularities among objects, like co-occurrence of particular features.

In order to observe the regularities that can be discovered by a self-organizing map, the study stimuli from the feature frequency experiment was fed to a Kohonen network using Viscovery SOMine software[3]. A map with 7 nodes was trained with tension parameter set to 0.5. The tension parameter determines the amount of fit to the data. If the tension is low, category representations become precise: nodes align better to the patterns in the data. If the tension is high, category representations are loose; nodes are more homogeneously aligned, and they might be further away from the actual data values. The tension parameter adjusts the balance between representation of individuals and representation of categories. If it is too high, categories become too general. If it is too low, the category information is lost. The resulting map is displayed in Figure 14-b.

---

[3] Viscovery SOMine is a data mining tool produced by Viscovery Software GmbH.

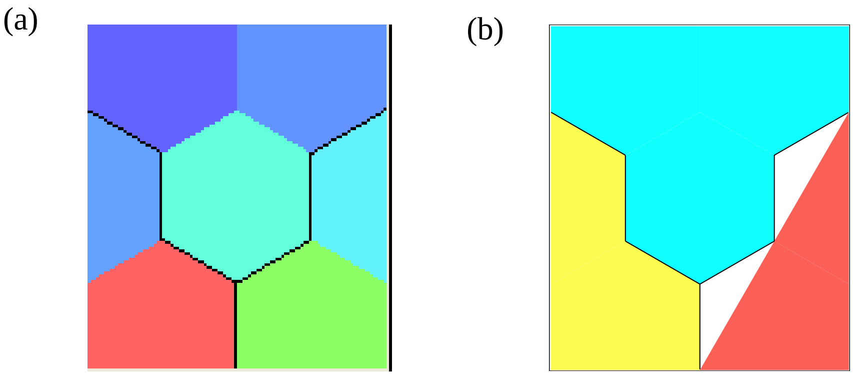


Figure 14 SOM clusters before and after training the SOM with 15 objects from the study phase of the feature frequency experiment, using Viscovery SOMine software. 7 nodes were used to construct the map. a) Initially, weights were assigned randomly to the map nodes. b) Three clusters formed after the training. Green area represents the segment sensitive to solid border and green color. Yellow area corresponds to the segment sensitive to oblique pattern and square shape. Red area represents the segment sensitive to mixed pattern and triangle shape.

The desired number of clusters can be set as a parameter. As the number of clusters increase, each cluster is tuned to more specific regularities. If the number of clusters is kept small, overlapping regularities can be kept in one cluster. For example, in the 3 cluster case, the cluster which is tuned to oblique pattern and square shape also contains a circle object. The reason for including an object without square shape is that it has a dashed border, and two other objects in the cluster also have a dashed border. In the extreme case, each cluster would represent an object in the study stimuli. Thus, there would be 15 clusters in the map, and each of them would be tuned to only one object.

Table 10. Results of the principal component analysis performed on the study stimuli.

	Component	
	1	2
border	.740	-.208
color	-.749	-.198
pattern	.472	.666
shape	-.435	.710

Finding regularities in the study stimuli is actually equivalent to finding the principle components of the stimuli. A principle components analysis was performed on the items in the study stimuli using SPSS. Each item was coded with a four digit number where each digit represented the corresponding feature value. After running the analysis, two components were discovered whose eigenvalues are greater than 1. The first component consists of border and color and the second component consists of pattern and shape. Table 10 displays the extracted components with factor loadings of each feature dimension. Even though there are other regularities in the study stimuli, only the color green-solid border pair and oblique pattern-square shape pair differ orthogonally. These pairs correspond to the FRFs in the behavioral experiments presented in Chapter 3.

After the formation of the clusters, test objects are presented. The clusters correspond to the output neurons of Layer 4 in VisNet which would respond maximally when the objects and correlations they were tuned to are presented. For each test object, each cluster will respond differently, according to the connection weights which were computed during the training. There will be competition among the clusters and the cluster with the highest value will win.

If it were only a categorization task, the mechanism of competition would be enough to determine the category of the object. However, the feature frequency experiment required participants to give an old/new response, in which participants should decide whether the value of the winning cluster is high enough to determine if the object was studied. For unstudied objects, there may not be any winning cluster, or the value of the winning cluster could be low, and this could be an indication of a new object.

There could be several different results after the training of the network with the study stimuli. Two of them will be discussed because of their significance in interpreting the results of the feature frequency experiment.

Each cluster representing only one object

In this setting, each cluster is tuned to only one object in the study stimuli. Assume that a new object whose three features are equal to an object in the study phase is presented. Each cluster will evaluate the stimulus and give a response. There would be competition among the clusters, and the response of the clusters which represent objects similar to the presented test object would be higher. If responses are equal, no one would win. If one cluster responds more strongly than the others, it would be the winner. However, a winning cluster would not necessarily mean that the object was studied. There should be a threshold-like mechanism for the old/new decision.

In the case of feature frequency variations among the study stimuli, unstudied test objects which have the frequently repeated features will be matched to a higher number of clusters than other test objects and there will be competition among all the responding clusters. However, there is no reason for why this strong competition would increase old responses for such objects. On the other hand, studied objects will be easily identified, since the cluster tuned to that object would respond maximally.

While this case predicts the high number of old responses for studied objects, it cannot account for the findings of the feature frequency experiment.

Clusters tuned to regularities in the study stimuli

If the trained map has a few clusters which are tuned to the regularities in the study stimuli, then there will be no separate representation for each study object. Clusters will represent a group of objects which are correlated along certain feature dimensions, like color and border. When a studied test object is presented, the cluster which represented the object during the study phase would respond maximally. For unstudied objects, if the object has features which match the regularities discovered by the clusters, these clusters will again respond maximally. If the object does not have any of these features, the clusters would respond much less. This mechanism is compatible with the results of the feature frequency experiment. The regularities discovered by a self-organizing map were green color-solid border and oblique pattern-square shape pairs. Unstudied objects with these features were significantly more frequently classified as "old" than unstudied objects without these features.

Thus, such mechanism is compatible with the results of the behavioral experiments.

Validity of the VisNet model with respect to the results of the feature frequency experiment

An analysis of the VisNet model using the stimuli from the feature frequency experiment indicated that a self-organizing network forming clusters tuned to the co-occurring features in the study stimuli can produce similar pattern of results. VisNet is selected as one of the convergence models and its more general properties were examined rather than the specific learning methods it adopts. For example, the trace learning rule is required by VisNet to produce viewpoint and translation invariant representations of objects. However, the trace learning rule is not appropriate to process successive presentations of different objects, since it regards all successive presentations as images from different viewpoints of the same object. The general self-organizing network structure of VisNet was more useful to represent frequent co-occurrences of particular features among the whole stimulus set, like the green color-solid border pair.

4.3.2 Item-matching models

Exemplar and prototype models constitute item-matching models. Their properties are explained in the literature survey chapter. Here, experimental results of the current study will be interpreted from these models' point of view. A detailed quantitative analysis of the prototype and exemplar models is presented in Chapter 6 and their performances with respect to the experimental findings are compared with the proposed model of the present study. In this section, the correspondence between the experimental framework of the present study and these models will be explained.

Exemplar Models

In the simplest exemplar model of categorization, each object would have a separate representation in memory (R. M. Nosofsky et al., 2011). With 15 training object in the feature frequency experiment, there would be 15 distinct object representations in memory. During the test phase, each test object would be compared to all 15 representations. If the object does not match one of these representations, it would be classified as new.

In the feature frequency experiment, participants frequently evaluated the oldness of the object wrong. In the simple exemplar model, evaluating a studied object as “new” can be a result of a weak or missing representation (because of lack of attention, etc) so that a match cannot be found for the incoming test object. On the other hand, evaluating an unstudied object as “old” is harder to explain. There would be no existing representation in memory for the incoming test object, and the memory system would still return a match for the object. Actually, the former (misses) was not very common among the responses; rather, false alarms rates were high. Thus, people were more inclined to give “old” responses for unstudied objects. In terms of the simple exemplar model, this means that memory representations are formed for all objects in the training phase, but the matching criteria is quite loose that a match is found for most of the objects.

What can be the effect of repeated features in such a setting? In the feature frequency experiment, high frequency of particular features during the study phase affected responses during the test phase. Unstudied objects with less frequent features had significantly lower “old” response percentage than unstudied objects with more frequent features. A simple matching mechanism is not sufficient to explain this feature frequency effect.

An exemplar model with a similarity measure

The GCM adopts the multidimensional scaling approach to determine similarity of objects in the feature space. In this approach, each object is a point in the feature space whose coordinates are values of features in different dimensions, like color, size, shape, etc. Similarity of two objects is determined by their distance in the feature space.

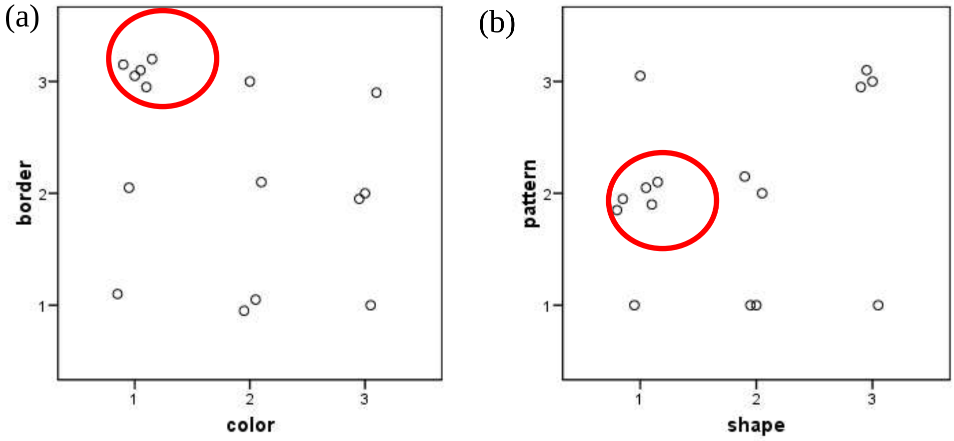


Figure 15 Objects plotted in two-dimensional feature space for the feature frequency experiment of the current study a) border-color b) pattern-shape.

In such a representational framework, study objects in the feature frequency experiment become points in the feature space defined by color, shape, pattern and border (see Figure 15). Since the features in the feature frequency experiment are discrete, calculating the distance between objects corresponds to checking for equal feature values for each feature/dimension. For example, consider the following objects:

O1: (solid, red, dotted, square)

O2: (dashed, blue, vertical, circle)

O3: (dashed, red, vertical, circle)

In order to calculate similarity of O3 to O1 and O2, feature values should be examined in four dimensions. For O1 and O3, only the values in the color dimension (red) are equal. On the other hand, O3 and O2 have equal values in three dimensions: border, pattern and square. Thus, similarity of O3 to O2 is higher than the similarity of O3 to O1, according to the multidimensional scaling approach.

When the participants are required to give old/new responses, an exemplar model with such a similarity measure would compute the similarity of each test item to each of the training items. There are two possible computations to determine whether the test object was studied or unstudied using the similarity measure. The first one is to find the most similar object in the training set and if the similarity is sufficiently high then give as “old” response. The second one is to sum all similarity values to obtain an overall similarity score (Nosofsky et al., 2011). The old/new decision then becomes a probabilistic decision depending on this similarity score. The higher the similarity score, the larger the probability of finding the test object “old” is. Nosofsky et al. note that his mechanism is an instance of a more general class of computations called “global matching models”.

If the test object has frequently repeated features, it will be similar to many study objects whose coordinates are equal to each other in two dimensions: either color and border, or pattern and shape. Thus, since the familiarity measure is the sum of all similarities, as suggested by Nosofsky et al., the sum for objects with frequently repeated features will be relatively high. Then, what happens if an object is highly similar to one object and not quite similar to the others? In the feature frequency experiment, studied objects were found “old” even when they were only similar to one object in the study phase, which is themselves. Thus, it seems that there is more to the old/new decision than just a summation of similarities. The number of similar objects does not seem to be a major factor, but the important thing is there should be objects either highly similar to the test object, or objects which are moderately similar to both the test object and to each other. It is like an attractor point in the feature space, either a very strong individual item or a neighborhood of items which pull together. If none of them is present, the object would be classified as “new”.

An actual implementation of an exemplar model with the stimuli from the experiments of the present study is presented in Chapter 6. The results are statistically analyzed and the findings are discussed in detail.

Prototype models

In the prototype model of categorization, object representations cluster around the prototypes of categories. In the feature frequency experiment, study items cluster around the two pairs of features: color green & black border, and square shape & oblique pattern, as can be seen in Figure 15. This clustering results in two prototypes. There are two options for a prototype: it can either be one of the objects belonging to the category, or an average of the existing objects. Since the feature values in the experiments of the present study are discrete, an average feature value corresponds to concatenations of feature values. For example, the first prototype becomes an object with green color, black border, shingle/dotted pattern and circle shape. Here, the pattern can be either shingle or dotted, since both values are at the center of the category. Objects with the remaining pattern, which is the diagonal line pattern, are relatively more distant from the center of the category. The second prototype was an item with square shape, oblique pattern, dashed/light border and blue color. Similarly, the border feature here has two values, since they are both in the center of the related category.

In the prototype models of categorization, the membership of an object to a category is determined by the distance of the object to the prototype of the category. For example, in Casale and Ashby (2008), the probability of an item to be classified into category A is the probability of the distance of the item to the prototype of the category B to be greater than the distance of the item to the prototype of category A:

$$P_{(\mathrm{A,B})}(\mathrm{A}|\mathrm{x}) = P(D_{xB} - D_{xA} > \varepsilon)$$

Here, $P_{(A,B)}(A|x)$ is the probability of item x to be classified into category A given two categories A and B. D_{xB} is the distance of item x to the prototype of B, and D_{xA} is the distance of item x to the prototype of A.

Using the above formula, the old/new decision can be modeled. This model assumes that object recognition uses prototypes instead of individual objects for making the decision. In the present study, the distance between items is defined as $DS(x_1, x_2)$ which is the discrete similarity of items x_1 and x_2. Thus, the probability of an object to be classified becomes:

$$P_{(old,new)}(old|x) = DS(x,p_C) = (\max DS(x,p_C)) - D_{xC}$$

where C is the category into which x was classified in the classification step, and p_C is the prototype of category C. The formula at the right side of the equality states that discrete similarity of an item to the prototype of a category is the reverse of the distance between the item and the prototype, thus the distance value is subtracted from the maximum possible value of the discrete similarity.

To compute the distance between the object and the prototype, the discrete similarity between the item and the category prototypes should be computed for every item in the test phase (18 items). The discrete similarity values would range from 0 to 4, where 0 indicates no common features with the prototype and the item, and 4 indicates an exact match between the item and the prototypical item. The item is then classified into the category which returned the highest similarity value to the prototype. The resulting similarity value can be used as the typicality of the object as a member of that category.

In such a setting, the frequency of features during the study phase directly affects category formation. The higher the repetition frequency of the feature is, the closer the prototype is to this feature. Thus, prototypes move towards the repeated features. In turn, test objects with frequently repeated features will be the most typical items of the two categories. Objects with only one of the frequently repeated features will be less typical, whereas objects without the frequently repeated features will be the least typical objects of the categories. The effect of this typicality on old/new responses is the main concern in such a model. The increase in “old” responses with the number of FRFs on the object can be explained with item typicality. However, this effect was seen only for unstudied objects. For studied objects, there was no effect of the number of FRFs on old/new responses. Prototype models do not have a property to explain such distinction between studied and unstudied objects. Studied objects are not necessarily close to the prototypes. They can even be very distant from the category center. However, the old/new response is totally unrelated to this distance.

An actual implementation of a prototype model with the stimuli from the experiments of the present study is presented in Chapter 6. The results are statistically analyzed and the findings are discussed in detail.

4.3.3 Feature models

The third model to be analyzed is the Theory of Event Coding (TEC) by Hommel and Colzato (2009). Since it is a dual process model, with a component for representing regularities (conjunction detectors) and another component for representing individual objects, it will be possible to discuss both mechanisms of binding, and see whether such a dual process model can explain the phenomena in the feature frequency experiment.

The TEC

According to the TEC model of Hommel et al. (2001), representation of objects in long-term memory corresponds to conjunction detectors which are formed through experience with several objects. In the feature frequency experiment, 15 objects were presented to the participants during the study phase. According to the TEC, each presentation would cause a temporary binding of object features in object files. Since two pairs of features co-occurred more frequently than other feature pairs, conjunction detectors would form for these feature pairs. Thus, in Hommel's notation, there would be two conjunction detectors in long-term memory as a representation of study objects: <green, solid> and <oblique, square>. These conjunction detectors would increase the strength of the bindings whenever objects with these feature pairs are presented.

In the old/new recognition task, first, study objects are presented. In the second part, test objects are presented and the participant is required to decide whether the object is studied or unstudied. Since an object file has a very short duration (4 s), the object files formed during the study phase are not expected to have any effect on the test objects.

Hommel states that conjunction detectors can store co-occurrences which are presented only once, but these representations would be weak since conjunction detectors would need several presentations of the feature co-occurrences. Thus, according to the TEC, representations of individual objects formed during the study phase should be weak and would not have much influence on the recognition of test objects.

During the test phase, 18 objects are presented one by one. When the first object is presented, a temporary object file is constructed. If the features of this object activate a conjunction detector, the features defined in the conjunction detector would reinforce the corresponding features. The first object in the test phase was a studied object. It also had a solid border which is a feature in the conjunction detector <green, solid>. Thus, it would also be supported by the top-down effect of a long-term representation. According to the TEC, the representation of this object in long-term memory would be weak, since the features of this object co-occurred only once. Besides, the conjunction filter expects color green when the solid border is perceived. Thus, TEC predicts that the percentage of "old" responses for this object would be low. However, the object had very high percentage of "old" responses.

The second test object is a studied object and has both oblique pattern and square shape, so it activates the second conjunction detector. Since the first object had no common features with the second object, their object files would neither conflict nor support each other. Thus, the percentage of "old" responses for the object should be high, and actually it is. The third test object displays similar properties, so it will be skipped in this discussion. The fourth object is worth examining since it seems to show some binding effects. It is an object which activates conjunction detector <green, solid> maximally, and conjunction detector <mixed, triangle> partially. On the other hand, the object file of the previous (third) test object conflicts with the object file of the fourth test object, since they both have mixed pattern. The activated conjunction detector <mixed, triangle> reinforces the object file of the third object further, and emphasizes the triangle feature. Thus, TEC predicts that the percentage of "new" responses for the fourth object would be high, and this was actually confirmed by the responses of the participants.

In summary, Hommel and Colzato (2009) consider the effects of two types of representations in their model: long-term representations of regularities (conjunction detectors) and short-term representations of individual objects. Even though both effects were observed in the feature frequency experiment of the present study, there was a third effect, the long-term effect of individual objects. Their model could not explain how studied objects could be identified as "old" significantly better than the unstudied objects even though unstudied objects were very similar to the studied objects. Thus, a model of old/new recognition should consider long-term memory representations of individual objects, in addition to the long-term representations of regularities.

4.4 Discussion of the results of the model analysis

Three representative models were analyzed with respect to the findings of the feature frequency and similarity experiments. A number of issues which are crucial to the understanding of underlying mechanisms were identified. First, the discovery of regularities in the stimuli by the convergence models is similar to extracting principle components of the stimuli. This is an efficient method of memory formation, where redundant information is omitted and only the frequently observed, possibly important, features are stored. The extraction of components is not as strict as keeping only the most relevant components, since there are many output nodes which are sensitive to smaller correlations but not very discriminative. Keeping the extraction criteria loose might be a solution for balancing the representations of generalizations and representations of individuals.

Second, there are many commonalities between the convergence models and item-matching models. Both types of models compute a value for the match between the test object and stored representations. Then if this value passes a certain threshold, the object is classified as "old". The computation in the convergence models is in the form of a competition among the output neurons, which is actually similar to the calculations in the item-matching models where either exemplars or prototypes compete to be the candidate match to the incoming test object.

Hommel and Colzato (2009) claim that convergence is not enough to explain connections among features since there would be a combinatorial explosion if all concurrently active features combine together. Also, they mention that it would require extensive training to learn these combinations. Beside, the convergence approach is based on the assumption that there should be correlations among features of the same object and dissimilarity of features of different objects so that distinct object representations can be formed. Rolls and Milward (2000) deny the need for extensive training, and argues that learning particular combinations is enough to recognize novel combinations. However, he agrees that the features belonging to the same object should be correlated. Hommel and Colzato add that even though features of real-objects are correlated, people are also capable of representing arbitrary combinations of features, as demonstrated by the performance in experiments. Thus, he suggests that a good model of object perception should both represent regularities and also arbitrary connections of features.

The cumulative evidence from the analysis of these three models shows that any model of the formation of visual object representations should account for the effects of regularities in the study stimuli on responses of the participants to the test objects. In this study, these regularities concern the frequency and the similarity of the objects. The model should be able to integrate regularities not only in subsequently presented objects, but also over extended periods of time. Thus, these regularities should be stored long-term.

The analysis of contemporary models in the field of object recognition and categorization revealed an important aspect of the findings of the feature frequency experiment. Convergence models assume that the convergence of low-level features to more abstract features occurs for building representations of objects which are viewpoint and translation invariant. On the other hand, models in the categorization literature assume that regularities are encoded to represent categories of object. However, the feature frequency experiment required neither categorization nor recognition from different viewpoints. Thus, it was shown that the encoding of regularities is not task specific; it is highly automatic and they are discovered even when there is no explicit requirement to do so. Thus, it is highly probable that both formation of viewpoint invariant object representations and category representations are results of this highly automatic and implicit process.

4.5 Summary of the Chapter

The main purpose of this chapter was to examine the correspondence between results of the experiments of the present study and existing findings of the false memory literature and models/theories of object recognition and categorization, especially in terms of the representational structures they employ. The findings from the false memory literature, especially the studies on category effects on false recognition, were closely related to the findings of the present study. Even though the experimental task did not require extraction of regularities in the study stimuli, the results were similar to the findings of Smith et al. (2000) where the same increase in false alarms was observed. Thus, their findings provided insights into the findings of the present study, pointing to the effects of categories, like the effect of graded category structure from the prototype theory, including the effects of typicality and output dominance.

While the results are compatible with the findings in the literature and principles of the existing models, no single theory/model can explain the phenomena observed in the experimental results of the current study. For example, the Generalized Context Model by Nosofsky et al. predicts that the sum of similarities would affect the recognition decision. According to the findings of the present study, the sum of similarities really affected old/new recognition, but it was not the only factor. The exact match of the test objects to the studied objects was also a major factor in addition to the sum of similarities.

The analysis of different models in the object recognition and categorization literatures pointed out a fundamental issue in building representations: the trade-off between representing individuals and representing regularities. If the representations are built only for individual items, the effects of regularities (features that appear multiple times as each object is introduced) cannot be demonstrated. If only regularities are stored, then the recognition of individual items becomes problematic. This is observed in Roll's VisNet where the parameters of the network adjust whether representations will be coarse or fine. Fine-tuned representations can recognize individual objects better, but they are less successful in identifying objects from different viewpoints. In the Theory of Event Coding (TEC) of Hommel et al. (2001), long-term representations are formed only for regularities. Representations of individual objects are temporary bindings of features. As a result, this theory cannot explain the successful recognition of studied objects in both feature frequency and similarity experiments of the present study, where the objects were presented only once during the study phase. The exemplar and prototype models are also good examples for the individuals/regularities dilemma. Exemplar models favor representations of individuals whereas prototype models favor representations of regularities.

An examination of the prototype and exemplar models with stimuli from the experiments of the present study showed that it is possible to make quantitative predictions about the processes which take place during categorization and recognition using the framework offered by these models. Since category typicality effects were shown to be a major factor affecting false memories, prototype models become especially important for providing a quantitative framework. Thus, the prototype and exemplar models will be analyzed quantitatively in Chapter 6 and their performance on the old/new recognition task will be compared to performances of human participants and the proposed model of this study.

As a result, an important constraint for a model of the formation of visual object representations is the capability to build representations for both individuals and regularities. More specifically, the model should explain various sides of the phenomena observed in the experiments of the present study: invariance to central tendencies for studied items and increase in "old" response percentages for unstudied items as the test item gets closer to the central tendencies. The next chapter presents the proposed model of the present study, developed on the Convergence-Divergence Zone Framework (Damasio, 1989), which is a general-purpose framework for the formation of memory representations. An evaluation of the proposed model according to the constraints mentioned above is presented in Chapter 6.

CHAPTER 5

CDZ-VIS: A MODEL OF VISUAL OBJECT REPRESENTATIONS AND CATEGORIZATION

After discussing psychological and neural models of visual perception, memory, and categorization in the previous chapter and evaluating them against the two present behavioral studies, a mechanism of the formation and activation of visual object representations and categories is presented in this chapter. The model, called “Convergence-Divergence Zone-Visual” (CDZ-VIS), is built upon the Converge-Divergence Zone (CDZ) Framework developed by Damasio (1989). While the CDZ Framework explains the general organization of memory representations, the CDZ-VIS Model focuses on the formation and activation of visual object representations, and incorporates specific hypotheses about category effects on object recognition. These hypotheses are based on findings from recent neuroscience literature. It will be shown that the CDZ-VIS model can account for the false memory effects observed in the similarity and feature frequency experiments.

5.1 The CDZ Framework

The convergence-divergence zone (CDZ) framework has been proposed by Damasio (1989) for explaining the structure of representations and the mechanisms of recall and recognition in general. It does not have a separate component for explaining visual object representations.

The reason for the selection of the CDZ Framework as a basis for modeling category effects on object recognition is that it already has a proposed mechanism for the formation of memory representations. The proposed mechanism has been shown by Damasio (1989) to be supported by numerous pieces of evidence from the neuroscience literature.

Alternatives to the CDZ Framework could be artificial neural network frameworks like the PDP proposed by Rumelhart and Zipser (1985). Such frameworks aim to model the processes responsible for various psychological phenomena. However, the mechanisms underlying the processes are not necessarily biologically plausible. Even though the units in these frameworks are models of neurons, they have additional claims which are not supported by current neuroscience research, like backpropagation. The CDZ Framework has been developed by Damasio to embody current findings on neural systems to serve as a basis for modeling cognitive phenomena. While ANNs enable modeling cognitive processes, CDZ Framework enables modeling both processes and underlying neural structures in a biologically plausible way.

The core principles of the CDZ framework on which the CDZ-VIS model was built are as follows:

- Features, objects and categories are represented by a hierarchical network of neurons, which extends from low-level visual areas to high-level integration areas.

- When a visual stimulus is presented, feature-representing neural populations start firing synchronously and connect to the neural populations in the associative layers, namely Convergence-Divergence Zones (CDZs). These CDZs correspond to object representations. Similar objects have closer CDZ representations.
- If an object is to be remembered, the episodic layer CDZs start firing, and they cause their lower layer CDZs to fire.
- Presentation of a stimulus a second time causes bottom-up firing of the object CDZ which represents this stimulus.

The basic property of this framework is the formation of uni-modal representations in sensory and motor cortices and through convergence to higher level regions, and integration of the uni-modal representations into multi-modal representations. This integration does not result from executive functions of higher levels, but from co-activation of the uni-modal representations. Damasio calls this “time-locked co-activation” to indicate that the co-activation takes places in a definite time frame. Instead of “co-activation”, “synchronous activation” will be used throughout this chapter.

The term “convergence zone” has been coined by Damasio to refer to the neural ensembles binding the feature-based fragments. The fragments of an entity or an event connect to a convergence zone in a many-to-one relation (see Figure 16). This bottom-up process is captured by the “convergence” part of the CDZ framework. When a convergence zone fires, all the sensory regions connected to that convergence zone become active. Thus, reactivation of sensory areas is possible through convergence zones in a one-to-many relation. If one of the fragments is reactivated, the convergence zone fires and the feedback connections cause reactivation of

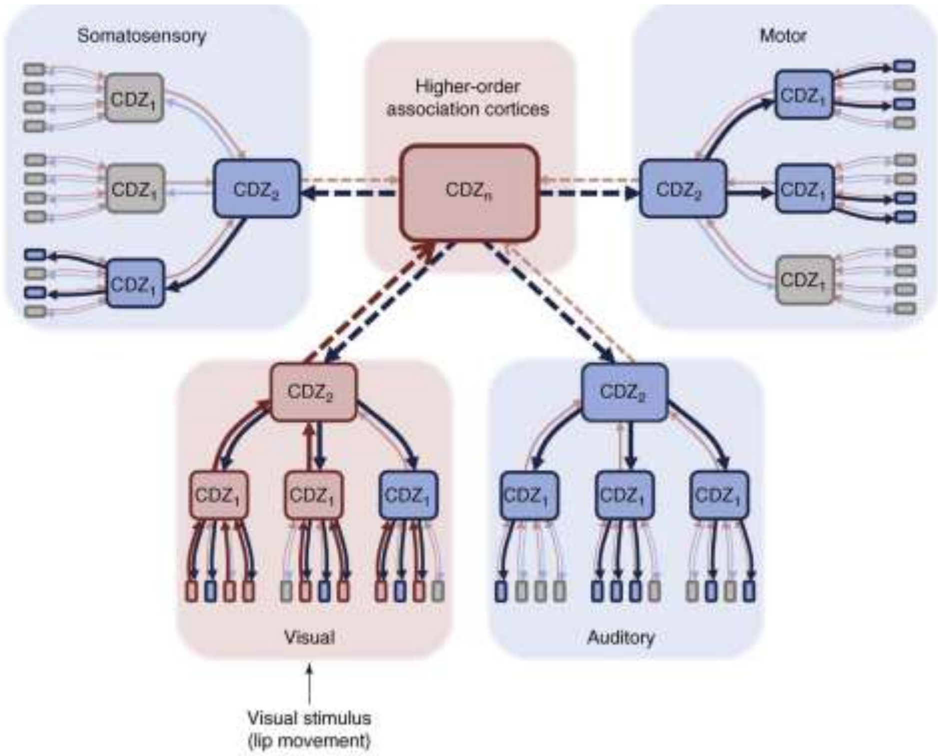


Figure 16. The CDZ framework (Meyer & Damasio, 2009).

all the other connected fragments. This top-down process is captured by the "divergence" part of the CDZ framework.

The convergence zone structure is crucial for explaining the formation of visual object representations and object categories in the CDZ-VIS model. An advantage of the convergence zone structures is that they are not just ad-hoc structures. Evidence of neural substrates for the convergence zones has been presented elsewhere (Damasio, 1989; Meyer & Damasio, 2009). Constraints determine the location of the convergence zones: domain of stimuli, number of components in an event and the anatomical design. A convergence zone develops when multiple regions are concurrently active. It has a threshold and levels of response, like low, moderate and high.

This framework explains behavioral phenomena using neurobiological constraints. Thus, it has both structured interpretation and dynamic realization (Barsalou, 2003).

The CDZ framework brings a solution also to the problem of process and representation dissociation. The representations of objects act like processors themselves. There is no separate process that operates on the features. Synchronous activation of individual feature representations converges to higher-level object representations. Actually, the definition of a system with representational structures and processes operating on these structures is not useful for neural systems. On the other hand, one can still separate the representational structure in the CDZ framework from the processes of convergence and divergence, but this division is not necessary. However, the possibility of making this division is an advantage over the traditional connectionist models, where the representations of items in the network are not separable from the processes, since it enables us to discuss findings from other Cognitive Science studies which use representational structures.

5.2 The CDZ-VIS Hypothesis

The following general hypothesis (which is the third hypothesis of the present study) underlies the CDZ-VIS model to be proposed:

H3: False memory effects in object recognition result from the interaction of convergent and divergent activations at convergence-divergence zones in the visual system according to the principles stated below.

This hypothesis attributes the effects observed in object recognition such as the category, association, and frequency effects to the workings of the bottom-up and top-down activations described as principles of the CDZ-VIS model based on the CDZ framework. It comprises the following seven principles:

P1: Object-representing neural populations (CDZs) which are firing synchronously converge to the neural populations in the upper associative layer CDZs, which correspond to category representations. The more features the objects share, the closer the category CDZ to the object CDZs in terms of layers.

P2: The more object CDZs converge to a category-representing CDZ, the stronger it becomes (firing becomes easier, can inhibit other category-representing CDZs).

P3: At each layer, competition occurs among CDZs to send signals through their divergent connections to lower layer object CDZs. The stronger CDZ as defined in P2 wins. The winner CDZ can inhibit firing of other CDZs. If two CDZs have equal strength, they cannot inhibit each other, and both can activate their lower layer object CDZs.

P4: If an object is presented, it activates object-layer CDZs. **The amount of activation on the CDZ depends on the match between the presented object and the objects represented by the CDZs.** An exact match results in the maximum possible amount of activation on the CDZ. The amount of activation determines the degree of familiarity, and decreases as the degree of match decreases. However, familiarity is not a fine-grained measure: it is either high, moderate, or low.

P5: A partial match requires recollection of the specific features of the partially activated representations. However, if the connection between the features and the object representation is not strong enough, recollection may not be possible and the degree of familiarity determines the final recognition decision. Recollection strength of an object-representing CDZ depends on the number of times it is activated by the features that converge to it. Similarly, recollection strength of a category-representing CDZ depends on the number of times it is activated by the object-representing CDZs that converge to it.

P6: If a winner category-representing CDZ activates an object-representing CDZ so that its activation level becomes higher than the activation level of the object-representing CDZs activated by the presented stimulus, the features of the object-representing CDZ activated by the winner category-representing CDZ are recollected. Recollection of the features of a category-representing CDZ is usually much more successful than recollection of the features of a single item, since features of a category are repeated several times.

P7: The final recognition decision is based on the match between the recollected features and features of the presented stimulus. If the features cannot be recollected because of low strength, degree of familiarity is used.

According to Principle 1, the same mechanism of the formation of individual object representations also creates the representations of categories. These categories are organized in layers, depending on the number of features which make up the categories. Even though the similarity of items is important for the location of the category representation in the brain, whether a category is formed or not does not totally depend on the similarity of objects. The important thing is the synchronous firing of neural populations.

In the CDZ framework proposed by Damasio, the contextual representations are at the highest layer of the hierarchy. The category and object representations are at much lower layers, closer to the sensory areas. When a contextual representation is activated, this activation spreads from higher layers to lower layers, including the category and object areas. This is called divergence. Divergence causes multiple CDZs to become active at the same time.

According to Principle 3, when multiple CDZs become active at the same time, they compete for activating their connections at lower levels. The winning CDZ can activate its lower level connections, and can inhibit activations of other CDZs in its layer. For a CDZ to win the competition, it should have strong connections with the context layer. The strength of the connections depends on previous experience with the objects in this specific context, as stated in Principle 2. Since the CDZs in the category layers are determined by the co-occurrence of particular features in the study set, the more frequent these features, the stronger the connections of the category layer CDZs are with the context.

An increasing number of neuroscience studies demonstrate that the visual system has competitive representations (Beck & Kastner, 2009). Beck and Kastner state that evidence is accumulating that the competition among the representational structures is affected by both top-down and bottom-up activations, and they provide a range of findings which can be explained by these competitive structures. Similarly, Fiete, Senn, Wang, and Hahnloser (2010) report that competition occurs for synapse growth/total synaptic strength at both pre- and postsynaptic neurons.

5.3 The CDZ-VIS Model

The structures and mechanisms explained in the CDZ-VIS hypothesis were modeled by defining the CDZs as representational structures whose properties are influenced by convergence and divergence processes. Table 11 displays correspondence of the entities in the model and the CDZ framework/CDZ-VIS hypothesis. A formal definition of the model specifications is provided in this section.

Table 11. Aspects of the CDZ framework and CDZ-VIS principles and how they are included in the CDZ-VIS model, and which properties are omitted.

Aspects from the CDZ Framework				
Aspect	**In theory**	**In model**	**Included**	**Omitted**
Feature representation	Convergence-Divergence Zone (CDZ)	A discrete feature value (e.g. red)	Feature-based representation	CDZ structure which represents continuous feature values
Spatial organization of CDZs	Objects with similar features have closer CDZs in the brain	Layer index determined by discrete similarity	Distance between CDZs determines the CDZ they converge to	Spatial information
Convergence	Synchronous firing of neurons connect them to CDZs	CDZs with common features connect to upper-layer CDZs	Concurrent activations cause convergence	Synchronous firing conditions other than similarity
Episodic memory	Top-down activation of episodic CDZs activate lower category CDZs	Top-down activation of episodic CDZs activate lower category CDZs	An activated episodic CDZ leads to activated object & category CDZs	How episodic CDZs form

Table 11 (cont.)

Aspects from the principles used in the CDZ-VIS hypothesis				
Competition of CDZs	CDZ with most lower-layer connections wins	CDZ with the highest strength value wins	Number of connections	Connection structure
CDZ strength	The number of low-level CDZs converging to it	The number of low-level CDZs converging to it	The number of low-level CDZs converging to it	Spatial information
Winning	Winner activates its divergent connections	Activation values of CDZs in divergence matrix of winner are increased	Increase in activation of divergent connections	Connection structure
Recollection	Recollected features are compared with the features of the stimulus	Compute discrete similarity of the recollected item and stimulus	Comparison of recollected item and presented stimulus	Neural basis of the comparison
Old/new decision	Based on interaction of convergent and divergent activations	Probabilistic evaluation of familiarity and recollection functions	Effects of category-representing CDZs combined with bottom-up activations	-

5.3.1 Formal definition of the model

The CDZ-VIS model consists of CDZs which are created and activated as each visual stimulus is introduced. A CDZ has three properties: a divergence matrix, strength and layer number. The elements of the divergence matrix of a CDZ are CDZs that converge to this CDZ and the strength of a CDZ is determined by the number of elements in its divergence matrix (the number of items that converge to the CDZ).

Object-representing CDZs are simplified in the model by assigning default values to strength and layer number. In theory, there is no structural difference between a category-representing CDZ and an object-representing CDZ. The reason for the simplification is that discrete (not continuous) features are defined as input to the network. In theory, features are also represented by CDZs. There is evidence that even very low-level features are learnt by experience (Fahle & Poggio, 2002).

For a category-representing convergence-divergence zone CDZ_{ij}, the divergence matrix D_{CDZij} is a 1xn row matrix of objects, $CDZ_{on,}$ that converge to the CDZ:

$$D_{CDZij}= [CDZ_{o1}\ CDZ_{o2}, \dots\ CDZ_{on}]$$

Note that the CDZ does not keep explicit information about the category it represents. It acts like an index which connects items in the same category. As seen in the notation above, a category-representing CDZ, CDZ_{ij}, is identified by an index, which consists of a layer and an id number (i and j, respectively). The layer number, *i*, of a CDZ is determined by the discrete similarity DS of the items that converge to it:

$$i = DS(CDZ_{o1}, CDZ_{o2}, \dots, CDZ_{on})$$

whereas the id number, j, is computed by incrementing the id of the last CDZ created in the same layer. Its purpose is to discriminate the CDZs in the same layer. Theoretically, CDZs do not need such indexes to be identified: they are organized according to the spatial properties (proximity, neuron type) of the converging CDZs.

The strength ST of a CDZ is determined by the size of its divergence matrix (the number of items to which it diverges):

$$ST_{CDZij} = n$$

where n is the size of the 1 x n matrix $D_{CDZij.}$

In the simplified object-representing CDZ, the elements of the divergence matrix are features, instead of CDZs:

$$D_{CDZoi}= [f_1\ f_2\ \dots\ fn]$$

The layer number is not calculated for object-representing CDZs since every object-representing CDZ is formed by a fixed number of features. The strength of an object-representing CDZ is also constant, which is 1, since every object-representing CDZ represents only a *single* combination of features. If the features were not discrete but continuous, different combinations of features could be represented by the same object-representing CDZ (for example, different views of the same object). In the ideal case, an object representation is a category itself, whose members are combinations of features which converge to the same CDZ.

As each stimulus is introduced, a new object-representing CDZ is created. Since layer number and strength of an object-representing CDZ are constant, the only operation is to construct the divergence matrix of the CDZ by storing the stimulus features.

If the presented object shares some features with the previously presented objects, their representations will become activated, together with the newly created object-representing CDZ. To accomplish this, the features of the presented object are compared with the features in the divergence matrix of the object-representing CDZs. If there are common features, a new category-representing CDZ is created, whose properties (strength, divergence matrix) and index (layer and id number) are assigned according to the calculations explained above. If there already exists a category-representing CDZ which has objects in its divergence matrix sharing the same features, a new CDZ is not created, but the strength of the existing CDZ is increased.

Object and category representations can be invoked for different tasks. In the case of item recognition, the participant is required to identify items from a previous context, which might require instantiation of episodic memories. As described in the CDZ framework, the episodic representations are at the top of the hierarchy. Thus, recognition of an item presented in a previous context requires activation of the CDZ representing this context. This results in the activation of all other units in the divergence matrix of this CDZ. The activation of all the CDZs in the divergence matrix starts a competition to activate their own divergence matrices. The winner is determined by the strengths of the CDZs, which are equal to the size of their divergence matrices.

$$A(D_{CDZkm}) = a_{max} \quad \text{if and only if } CDZ_{km} \in D_{CDZij} \text{ and } k = \text{argmax}_k\, n$$

where n is the size of the divergence matrix of CDZ_{km}. In other words, the winner is the CDZ with the most divergent connections to the lower layers. The argmax function returns the value of index k which belongs to the CDZ with the largest divergence matrix. All the CDZs in the divergence matrix of the winner CDZ are activated whereas all the remaining CDZs in the current layer are inhibited. These patterns of activations will percolate from upper layers of the hierarchy towards lower layers, reaching to the object-representing CDZs. At the same time, the presentation of the object will cause bottom-up activation. The recognition decision is based on the interaction of the bottom-up and top-down activations.

The familiarity function is as follows:

$$F(S) = \begin{cases} f_{max} & \frac{DS(S, D_{CDZoi})}{n} = 1 \\ f_{moderate} & 0 < \frac{DS(S, D_{CDZoi})}{n} < 1 \\ 0 & \frac{DS(S, D_{CDZoi})}{n} = 0 \end{cases}$$

where S is a row matrix containing features of the presented object, D_{CDZoi} is the divergence matrix of CDZ_{oi} and n is the size of the two matrices, with the assumption that they will be of equal size. The familiarity function takes the presented stimulus as input and returns one of the three familiarity values. It compares the amount of similarity with the highest possible amount of similarity. If they are equal, the highest familiarity value f_{max} is returned. If the similarity between the presented object and the object-representing CDZ does not reach the maximum, but not 0 either, a moderate familiarity value is returned. The familiarity function does not distinguish between every level of activation: the levels between the maximum and 0 are considered as the same. Thus, the familiarity function does not provide a fine-grained measure for recognition. However, only when the familiarity function returns $f_{moderate}$, the recollection function provides further information about the match between presented stimulus S and recollected feature matrix C:

$$R(S,C) = \begin{cases} \frac{\sum_{j=0}^{n} \sum_{i=1}^{m} f(ci, sj)}{q} & \text{if} \quad q>0 \\ 1 & \text{if} \quad q=0 \end{cases}$$

where q is the size of C (number of successfully recollected features) and f(x,y) is a function which returns 1 if x=y and returns 0 otherwise. The recollection function checks if the presented stimulus includes recollected features and returns the ratio of included features to the total number of recollected features. If no features could be recollected, the function returns 1, making it ineffective.

The recognition decision is made by considering the results of both familiarity and recollection:

$$P(\text{“old”}) = \begin{cases} p_{max} & \text{if } F(S) = f_{max} \\ R(S,C)(p_{max}-p_{unit}) & \text{if } F(S) = f_{moderate} \\ p_{min} & \text{if } F(S) = 0 \end{cases}$$

where p_{max} is the highest possible probability of an “old” response in the system. If the presented item causes maximum familiarity, it is identified as “old” with high probability, without need for recollection. Studied items can result in maximum familiarity. If familiarity level is moderate, recollection is needed. The value returned from the recollection function is used as a multiplier to the moderate familiarity level. Finally, if the familiarity of the object is 0, the probability of an “old” response is at minimum. The actual values for maximum and minimum probabilities depend on the neurons

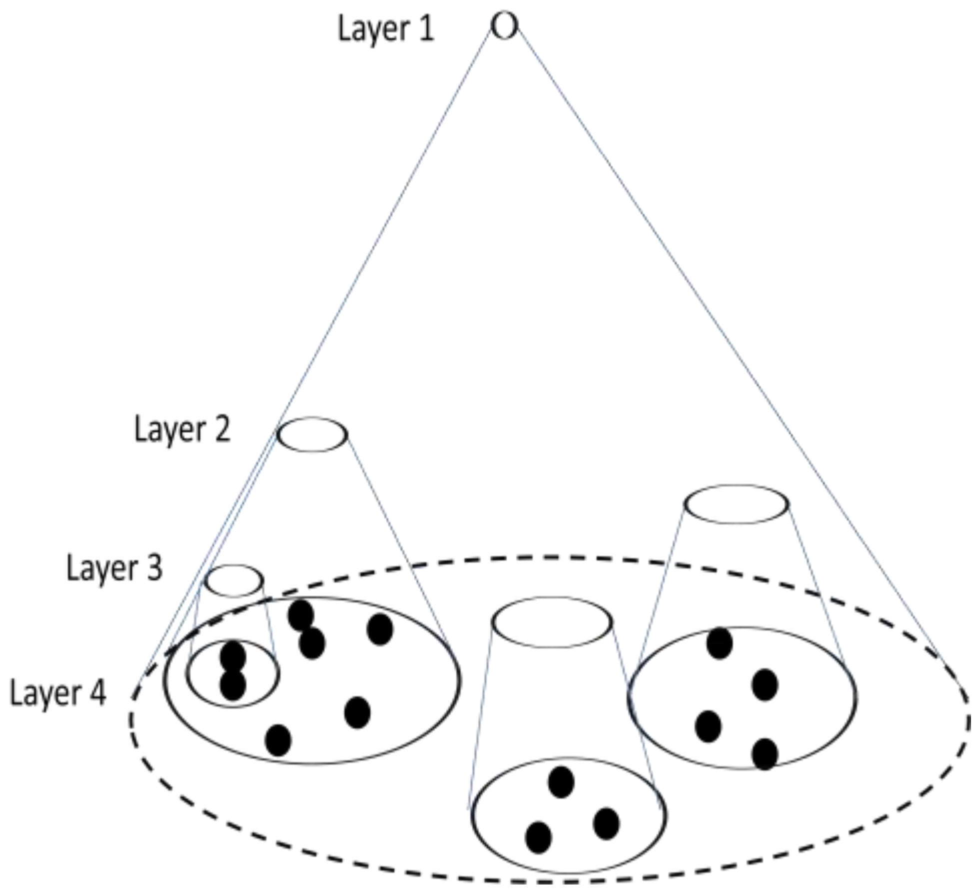


Figure 17. Depiction of the object and category representations at various layers in the CDZ-VIS model. Black dots correspond to object-representing CDZs. Circles at Layers 3, 2, and 1 correspond to category-representing CDZs. Feature-representing populations are omitted for the sake of clarity, but they appear at lower layers.

making up the CDZs. Any implementation of the CDZ-VIS model can employ existing models of neurons to determine the maximum and minimum values for the probabilities, as will be shown in the simulation section. Figure 17 demonstrates a sketchy model of the visual object and category representations implied by the CDZ-VIS model.

5.3.2 Simulation of memory formation in the CDZ-VIS Model

The model was simulated by implementing the model specifications in the C++ programming language. The code can be found in APPENDIX E.

The program reads items in the study stimuli from a text file one-by-one, creates an instance of the CDZ class, and stores the features of the object as the divergence matrix of this CDZ. The layer number of the CDZ is determined by the number of features it encodes. The layer of an object-representing CDZ is

$$S_{CDZij} = | D_{CDZij} | = 4$$

since it encodes the four features of the item (items in the behavioral experiments of this study consisted of the four features: color, border, pattern, shape). If the currently read item has common features with previously encoded items, a new CDZ instance is created, but at an upper layer. Since the layer information for the CDZ is determined by the number of features it represents, the number of common features, which is DS(object1, object2), determines the layer information.

For the formation of the hierarchical representation of the feature frequency stimuli, 18 objects were provided as input to the program in a text file. The resulting hierarchy is displayed in Figure 18. It is like a giant single representation which represents all objects in the study stimuli. As can be seen in the figure, individual object representations are only a subset of the overall hierarchy. There are many upper layer CDZs which represent different categories.

For the formation of the hierarchical representation of the similarity stimuli, 9 objects were provided as input to the program in a text file. The resulting hierarchy is displayed in Figure 19.

Layer 1

(11)x4 (12)x4 (4)x7 (5)x4 (9)x4 (7)x4 (3)x4 (2)x7
(1)x4 (10)x7 (6)x4 (8)x7

Layer 2

(4,9) (3,7) (4.2)x5 (2,10)x4 (2,8)x2 (6,11) (4,10)x2
(4,8) (8,10)x5 (9,11) (6,1)x2 (7,12)x2 (5,12) (8,12)
(5,8)x2 (1,8) (5,3)

Layer 3

(1,9,11) (4,9,11) (4,2,10) (4,2,9) (4,2,12) (4,2,7)
(2,8,10) (6,8,10) (5,8,10) (3,8,10)

Layer 4

(5,3,7,12) (4,1,9,11) (4,2,9,10) (6,3,7,11) (4,2,8,12)
(5,2,8,10) (6,1,9,11) (4,2,7,10) (4,3,8,10) (4,2,9,11)
(6,1,8,10) (4,2,7,12) (6,2,8,10) (5,1,8,12) (5,3,8,10)

Figure 18. The resulting hierarchy after the objects from the study stimuli of the feature frequency experiment were fed to the model. Numbers inside the brackets are codes for feature values. For example, (4,2,9,10) in Layer 4 is a CDZ, which represents an object with green color, black border, dotted pattern and square shape. Another example, (4) in Layer 1, is a CDZ which represents objects with green color. The number after the cross is the strength of the CDZ, which is equal to the number of times the feature was displayed during the study phase.

Layer 1

(4)x3 (7)x3 (8)x3 (5)x3 (9)x3 (6)x3

Layer 2

(1,12)x3 (2,10)x3 (3,11)x3

Layer 4

(4,2,9,10) (4,1,8,12) (4,3,7,11)
(5,1,7,12) (5,2,8,10) (5,3,9,11)
(6,1,9,12) (6,2,7,10) (6,3,8,11)

Figure 19. The resulting hierarchy after the objects from the study stimuli of the similarity experiment were fed to the model. Numbers inside the brackets are codes for feature values. For example, (4,2,9,10) in Layer 4 is a CDZ, which represents an object with green color, black border, dotted pattern and square shape. Another example, (4) in Layer 1, is a CDZ which represents objects with green color. The number after the cross is the strength of the CDZ, which is equal to the number of times the feature was displayed during the study phase.

5.3.3 Simulation of old/new recognition in the CDZ-VIS Model

The old-new recognition in the CDZ-VIS model was simulated using 15 test items from the feature frequency experiment and hierarchical memory representation produced in the first part. For each item, levels of activation $A(CDZ_i)$ were calculated for each object-representing CDZ in the hierarchical representation. Familiarity of the item was calculated using F(S) for CDZ_i with the highest activation value. If the familiarity was moderate, recollection was attempted, using function R(S). Probability of an "old" response was calculated by evaluating the recollection multiplier and level of familiarity. The probability scale in Verhoef, Kayaert, Franko, Vangeneugden, and Vogels (2008) was used where the highest normalized value was $a_{max}=0.8$, the minimum normalized value was $a_{min}=0.2$ and unit change in the normalized value as $a_{unit}=0.1$. The old/new decision was produced with the resulting probability. This 15-item process was repeated 20 times to simulate an experiment with 20 participants.

5.3.4 Comparison of the performance of the CDZ-VIS Model with human data

Feature frequency experiment

For the first category of FRFs (green color, black border), data from 20 human participants was compared with model-produced data using a repeated-measures ANOVA (# of FRFs: 0, 1, 2; study condition: studied, unstudied; source: human, model). Study condition and the # of FRFs were highly significant whereas there was no effect of source. (study condition: $F(1,38)= 75.86$, $p< 0.0001$, $\eta^2= 0.67$; # of FRFs: $F(2,76)=17.54$, $p<0.0001$, $\eta^2=0.32$; study condition * # of FRFs: $F(2,76)=8.43$, $p<0.0001$, $\eta^2=0.18$; oldness * # of FRFs * source: no interaction effect). Statistically, the results were the same for human and model participants (see Figure 20-a,b).

For the second category of FRFs (square shape, oblique pattern), data from 20 human participants were compared with model-produced data using repeated-measures ANOVA (# of FRFs: 0, 1, 2; study condition: studied, unstudied; source: human, model). Study condition and the # of FRFs were highly significant whereas there was no effect of source. (study condition:

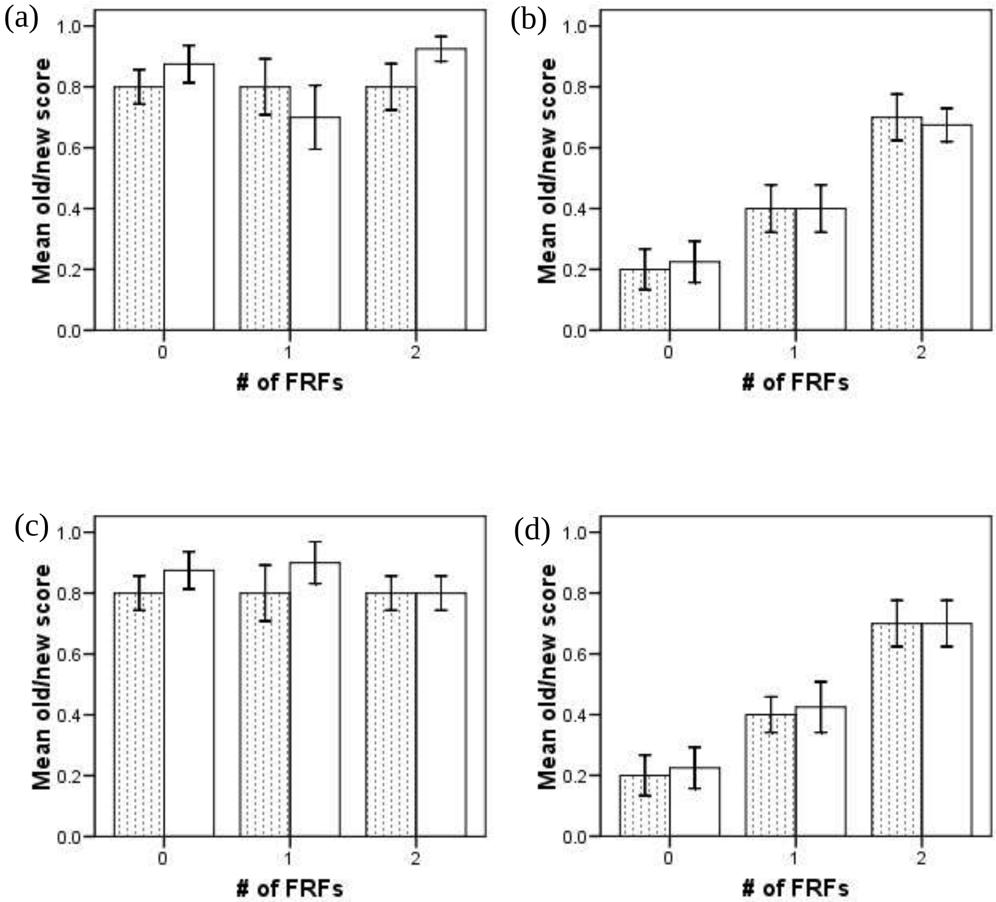


Figure 20. Comparison of model and human performance for the feature repetition experiment. Bars with dotted pattern indicate model performance. Error bars represent standard error of the mean. (a) Studied objects, first pair of FRFs (b) Unstudied objects, first pair of FRFs (c) Studied objects, second pair of FRFs (d) Unstudied objects, second pair of FRFs

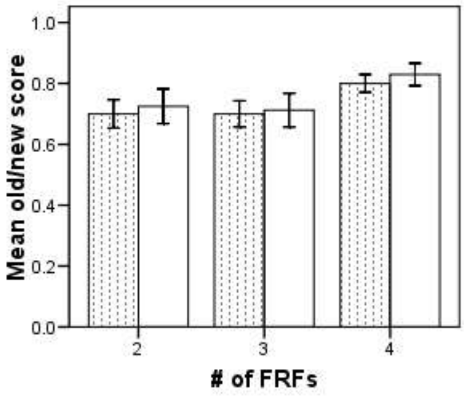


Figure 21. Comparison of model and human performance for the similarity experiment. Bars with dotted pattern indicate model performance. Error bars represent standard error of the mean. (a) Studied objects (b) Unstudied objects

$F(1,38)= 71.98$, $p< 0.0001$, $\eta^2= 0.65$; # of FRFs: $F(2,76)=10.73$, $p<0.0001$, $\eta^2=0.22$; study condition * # of FRFs: $F(2,76)=17.25$, $p<0.0001$, $\eta^2=0.31$; study condition * # of FRFs * source: no interaction effect). Statistically, the results were the same for human and model participants (see Figure 20-c,d).

Similarity experiment

Data from 20 human participants was compared with model-produced data using repeated-measures ANOVA (similarity: 2, 3, 4; source: human, model). $F(2,76)=4.24$, $p<0.05$, $\eta^2=0.1$; similarity * source: no interaction effect). Statistically, the results were the same for human and model participants (see Figure 21).

5.4 Limitations of the CDZ-VIS Model

There are three critical issues to be questioned about the CDZ-VIS model: Extensibility, generalizability, and robustness. Extensibility is the ability of the model to deal with large amounts of data beyond the experimentally demonstrated portion. Generalizability is the ability of the model to respond when the model is tested in different experimental settings, not only in the experimental setting of the present study. Robustness is the ability of the model to respond properly when either the components of the model or input to the model are damaged.

The CDZ-VIS model can be identified as extensible, since there is no parameter restricting the number of features or feature combinations to be processed. Whenever an object is presented, a new CDZ is created. The number of CDZs thus depends on the medium the model is implemented on. The model does not have a restriction itself.

In terms of generalizability, the model has some limitations, but these limitations can be overcome by extending the scope of the model to include lower-level feature representations. These limitations are mostly related to the discrete values of features. For example, the color feature had three values: red, green, and blue. Input from the environment is actually continuous. There are various intensities of red, green and blue. In order to overcome this limitation, there should be feature-layer CDZs, just one layer below the object-layer. There can be many feature layers below, until the rawest input is reached. This will enable the formation of feature categories, e.g. different intensities of red will converge to red.

Another limitation is about the similarity function. While comparing features of two CDZs, it requires feature values to be exactly equal. When there are subtle differences between the feature values, this similarity function would decide that two features are different. However, this may not be compatible with real world data. There will always be a change in features even when the same object is presented. Thus, a more plausible similarity function would employ fuzzy logic for comparison. Another limitation about the similarity function is that the number of features should be equal in the compared CDZs. This limitation can also be easily overcome by adding a mechanism to compare only the features whose values are available.

Another limitation is about episodic representations. CDZ-VIS does not model how episodic representations are formed, so the CDZs are assumed to receive activation from the episodic layer, and competition starts at each layer.

If the similarity among study items were not as high as in the experiments of the current study, less category-layer CDZs would form. In the case where no shared feature between study items existed, no category-layer CDZ would form. Then, recognition performance would depend on the match between features of the presented test item and features of the existing object representations. If the match is high, the test object can be falsely recognized as “old”. Thus, the CDZ-VIS model is generalizable also for stimuli sets involving different statistical properties.

Regarding the effects of damage to the representational structures, the CDZ-VIS model can be said to be very robust. If an object-representing CDZ is removed, and the object is presented again, another object-representing CDZ, with a divergence set similar to the features of the presented object can be activated. For example, if the representation of a specific green apple is removed, the representation of a red apple will be activated. Of course, the amount of activation will not be as high as the original CDZ: it will depend on the number of common features between the divergence sets of the CDZ representing the red apple and the CDZ representing the green apple. Whether the red apple will be identified as the green apple will depend on the recollection performance, which depends on the amount of previous exposure to these apples.

5.5 Summary and contribution of the modeling study

Important process-related aspects of the CDZ-VIS theory were modeled leaving details of the structural aspects aside. However, the level of abstraction in process modeling is still in the neural level. Table 10 displays aspects which are and which are not included in the model. The omitted properties are mostly related with the physical structure of the neural connections. Even though the physical structure is also important, the aim of this model is to show that the principles mentioned in the CDZ-VIS hypothesis are capable of explaining behavioral phenomena.

The model was simulated with computer software to demonstrate this capability. First, the feature frequency experiment was simulated by providing the model with the study stimuli and obtaining 20 model-participant data. Results were statistically compared to the results of behavioral experiments and there were no significant difference. The same process was repeated for the similarity experiment, and results were again not significantly different from human data. Thus, the same model could account for two different experiments, which had two different sets of stimuli.

The similarity and frequency variables indicate two different aspects of the CDZ-VIS model. Similarity of a test object to the study objects directly affects the familiarity of the test object. Indeed, the familiarity of an object is determined by the amount of similarity between the object and object representations that formed during the study phase. If there is an exact match between the features of the test object and the features of an object representation, the object becomes highly familiar. No match results in zero familiarity. If a number of features match, then a moderate level of familiarity occurs, and a recollection process is required. Feature frequencies have a totally different effect. They cause formation of category-layer CDZs, by activating several object-representing CDZs simultaneously, causing them to converge to an upper-layer CDZ. Thus, repetition frequency causes top-down effects in object recognition, where as similarity of test objects to study objects cause bottom-up effects.

The similarity of study objects to each other is determined by the repetition frequency of features, therefore it is not separately analyzed. If a feature repeats several times, there will be many objects sharing this feature, becoming similar in terms of this feature. If multiple features repeat together, then the similarity among the objects sharing this feature combination becomes higher. Thus, similarity of study objects to each other is a result of the repetition frequency of the features, both individually and in combination.

The success of the model to simulate results of the feature frequency and similarity experiments is an indication of the plausibility of the mechanisms suggested by the theory. However, the analysis would not be complete until the model was shown to provide a better explanation than existing theories for the false memory phenomena. Therefore, the next section will compare the model's performance with other models in the same domain.

CHAPTER 6

VALIDATION OF THE CDZ-VIS MODEL

This chapter aims to show that the CDZ-VIS model can satisfy constraints imposed by the feature frequency and similarity experiments much better than existing models in the categorization literature. There are two main streams of explanations in the false memory literature: category effects and association effects. For category effects, the CDZ-VIS model will be quantitatively compared to two dominant models in the categorization literature: prototype and exemplar models. Finally, the CDZ-VIS model will be qualitatively compared to an artificial neural network model, the PDP by Rumelhart and Zipser (1985), which is an associative model (Rogers & McClelland, 2004).

6.1 The CDZ-VIS model and associative processes vs. category effects in false memory

An explanation of false memory phenomena with the CDZ-VIS theory involves both memory representations and decision processes. The structure of memory representations that form during the study phase is important in determining the activated representations during the test phase. During the test phase, these activations determine the decision making process. Actually, it is hard to differentiate decision process and representation, since the decision process is largely dependent on the activation of these representations.

According to the CDZ-VIS model, the recognition of an item is affected by a number of factors. A basic factor is the strength of the connections between the item representations and features. This factor affects the recollection performance during recognition. The stronger the connection between the features and the item, the more details about the item are recalled. Another important factor is the strength of connections between the item representation and episodic representation. Episodic representations are at the highest level in the representational hierarchy. In the experiments of the present study, episodic representations include the objects in the study phase and all the other information like the environmental setting, the experimenter, and other people around. The strength of connections between the item representation and episodic representation affects the familiarity of the item during recognition. However, familiarity does not just depend on encoding processes which takes place during the study phase. The amount of activation of the item representations during the test phase is a major factor affecting familiarity. Another factor is the formation of categorical representations when items are presented. This is similar to "extracting the gist" of a set of stimuli in the literature (Brainerd & Kingma, 1984; Brainerd & Wright, 2005). The CDZ-VIS theory does not make a distinction between a gist representation and a category representation. When neurons synchronously fire, the spread of activation from these neurons intersect at some region, called a CDZ, and connections form between these neurons and that CDZ. If a category or a gist forms during the presentation of items, it will be connected to the episodic representation. Thus, during recognition, its representation will also be activated, and will affect the recognition process. For example, in the feature frequency experiment, two categories formed during the study phase in the CDZ-VIS model. In the test phase, these categories affected responses for unstudied objects. Whereas the false memory literature seems to be divided into two - association-based theories and gist-based theories, the CDZ-VIS theory involves both association based and gist based processes in item recognition. However, it is not a hybrid theory. Whether a gist is extracted or an association is activated depends on the items in the studied list, not on the processes involved.

For example, in the “converging associates” phenomenon, items presented during the study phase activate representations of associated items and these items later affect recognition of test items. The associations do not necessarily depend on the similarity of the items; the crucial factor is that a connection between the items must be set previously. The stronger the association between the presented items and associated items, the stronger the effect on recognition during the test phase. There are two main explanations of this phenomenon. The first one is the spread of activations during encoding, which activates associated items, to the degree of previous association strength. The second explanation is the gist formation, which claims that the gist of the presented stimuli is extracted which later increases false alarms for the gist item. The gist explanation cannot account for the cases when there is no common gist of the presented items. It also cannot explain the effect of association strength. According to the CDZ-VIS theory, presentation of an item during the study phase causes spread of activation starting from low-level visual areas through object and category representations reaching to episodic representations. During the study phase, this spread of activation activates previously built connections, if there exists any, too. Since there were no previously built connections, other than the ones that were built during the study phase, in the experiments of the current study, such connections did not exist. However, if there were such connections, as in the converging associates procedure, they would be activated, together with the presented item. Thus, according to the main principle of the CDZ-VIS theory, which states that synchronously active regions converge to upper layer CDZs, both representations would converge to the same episodic representation, if any previous associations existed. Even though the CDZ-VIS theory and association-based theories both mention the spread of activation during encoding, they differ in explaining the effect of the activated representations. While association-based theories state that the effect is due to the residual activation on the critical lure, CDZ-VIS theory predicts that the effect is due to the activation of the critical lure during the test phase because of its connection to the episodic representation in the higher layer CDZ. In that regard, the effect of the critical lure is similar to the effect of the gist (or category) representation, which affects item recognition during retrieval. Thus, the effect of the gist and the effect of the critical lure are similar in the CDZ-VIS theory, since they connect to the episodic representation during encoding, and both affect false memory during retrieval. However, the gist may have a smaller effect if the items presented during encoding do not activate it strongly enough during the encoding, which depends on the strength of the connection between the item and the category.

6.2 The CDZ-VIS model and category effects on false recognition

The two explanations in the false memory literature, associative processes and category effects will be examined separately. In this section, two fundamental models of the category formation be examined and compared to the CDZ-VIS model quantitatively through a simulation of their behaviors in the feature frequency and similarity experiments. In section 6.3, the CDZ-VIS model will be discussed with respect to Rogers and McClelland's (2004) version of the PDP model.

6.2.1 The relationship between the prototype theory and CDZ-VIS

To test the effect of category output dominance, Smith et al. (2000) used category lists by omitting the most typical item determined by the typicality ratings in Rosch’s (1975) original lists, and later requiring the participants to recall the items in the list. The category name was displayed before the category members were presented. A between-subjects “delay” variable was also tested (recall was requested either after a single list or after every category list had been presented). They obtained the same result as in the current similarity experiment where there was a significant difference between studied and unstudied objects. In addition, they also found that there was a correlation between output dominance and frequency of false memory. This correlation was not significant when performed only for critical intrusions. Since the critical intrusions were determined by typicality, it can be inferred that output dominance is a more important factor for false alarms than typicality. Actually, similar results were found in the experiments of the current study. In the similarity experiment, the similarity of the items corresponds to how typical the test items of the study set were. In line with the findings of Smith, a significant difference was found between “old” responses for studied and unstudied items, but no significant difference was found for levels of similarity. On the other hand, there was a highly significant difference between levels of the FRF variable in the feature frequency experiment. As the number of the FRFs on the item increased, the probability of an “old” response for the item increased.

In the CDZ-VIS model, the number of false alarms depends on the discrete similarity of the test items to the study items and the strength of the categories that formed or activated during the study phase. The strength of a category depends on the number of items that belong to the category. In such a setting, the typicality of an item is the amount of activation in the CDZ which represents this item when the category-representing CDZ is activated. On the other hand, output dominance is the probability of an item to be activated maximally when the category-representing CDZ is activated. They both result from the divergent connections from the category-representing CDZ to the item-representing CDZs. The amount of activation on an item representation depends on the distance of the item representation to the item-representing CDZs activated by the category-representing CDZ through the divergent connections, which is affected by the similarity of the item to the items in this target region. Such an explanation of typicality is perfectly compatible with the typicality definition of Schmidt, who states that typicality of an item is the similarity of the item to the "conceptual core" of the category. On the other hand, output dominance is more related with the strength of the divergent connections, which is determined by the presentation frequency of the item. If an item is seen several times, its connections with the category will be stronger than the connections of other members of the category with the category. Even though they seem to be equivalent in meaning, they correspond to different properties of the network. This property is compatible with Smith's finding that typicality and output dominance were actually correlated in their experiment.

In the third experiment, Smith et al. seek a linear relationship between category output dominance and false recall. There were three levels of the category output dominance variable: low, medium and high. They primed each of these levels with an additional task before the experiment in which they required participants to rate the pleasantness of the words. For example, an item with low output dominance was presented to the participant. The participant was required to rate the pleasantness of the item. This way, the category to which the item belonged was primed (the category was expected to be slightly primed, since the item has low category output dominance).

In the item recognition with CDZ-VIS, the familiarity of the item determines the first step. If it is high enough, the object is found “old”. If the familiarity of the object is not high enough, the recollection step begins. During recollection, categories and associations can have an impact on the final decision. The effects caused by the categories depend on their strength, which is determined by the number of category members. A strong category can activate the representations of its members through divergent connections. The effect of output dominance takes place during this step: features of the representations of items with the highest output dominance will be recollected. Thus, only features of the test objects with the high output dominance will match with the recollected features, and will be found “old”.

Typicality can be correlated with output dominance in some situations because the repetition of objects causing output dominance can also change typicality of objects by moving the center of the category toward objects with high output dominance. Thus, it can be hard to differentiate the effects of typicality and output dominance, as in the feature frequency experiment, where similarity of the test objects to the study objects and repetition frequency of the features were correlated. A separate experiment testing only the effect of similarity revealed that similarity had not significant effect on old/new responses. In the CDZ-VIS model, typicality corresponds to the similarity of the test item to the center of the categories that form during the study phase. Since the output dominance determines whether the features of an object representation activated by a category representation will be recollected or not, the presented object will only be compared to the object representation with the highest category output dominance, regardless of its category typicality.

6.2.2 A quantitative comparison of the performance of CDZ-VIS, Exemplar and Prototype Models

To provide further evidence for the contribution of the CDZ-VIS model for explaining human performance in the feature frequency and similarity experiments, predictions of three models were compared quantitatively. Exemplar and prototype models were constructed according to the model definitions in the literature, which will be explained in detail below.

In the prototype model, a category is represented by a prototype which corresponds to the central tendency of the items in the category. A prototype can be a member of the category, or a more abstract construct like a feature bundle. For implementing the prototype model, the central tendency of the items in the study phase was calculated by finding the mode for each feature type. Color green, black border, square shape, and oblique pattern were the modes for each type. Second, the modes were calculated pairwise: color green-black border and square shape-oblique pattern were the modes for the pairs color-border and shape-pattern. For each pair, the modes for the other feature types were computed. Two prototypes were obtained as a result. The first prototype was an item with green color, black border, shingle/dotted pattern and circle shape. Two values were included for the pattern feature since there were two modes for that type. The second prototype was an item with square shape, oblique pattern, dashed/light border and blue color. Two values were included for the border feature since there were two modes for that type.

The membership of an item to a category was calculated according to the formula given by Casale and Ashby (2008):

$$P_{(\mathrm{A,B})}(\mathrm{A}|\mathrm{x}) = P(D_{xB} - D_{xA} > \varepsilon)$$

which states that the probability of an item to be classified into category A is the probability of the distance of the item to the prototype of the category

B to be greater than the distance of the item to the prototype of category A. Here, $P_{(A,B)}(A|x)$ is the probability of item x to be classified into category A given two categories A and B. D_{xB} is the distance of item x to the prototype of B, and D_{xA} is the distance of item x to the prototype of A.

The old/new decision was modeled using the distance between the item and the prototype. In the present study, the distance between items is defined as $DS(x_1, x_2)$ which is the discrete similarity of items x_1 and x_2.

$$P_{(\text{old,new})}(\text{old}|x) = DS(x,p_C) = (\max DS(x,p_C)) - D_{xC}$$

where C is the category into which x was classified in the classification step, and p_C is the prototype of category C. The formula at the right side of the equality states that discrete similarity of an item to the prototype of a category is the reverse of the distance between the item and the prototype, thus the distance value is subtracted from the maximum possible value of the discrete similarity.

For every item in the test phase (18 items), the discrete similarity between the item and the category prototypes were computed. Computed values ranged from 0 to 4, where 0 indicates no common features with the prototype and the item, and 4 indicates an exact match between the item and the prototypical item. The item was classified into the category which returned the highest similarity value to the prototype. As the result of the computation, a category and a typicality value were returned for each test item. The typicality values were scaled to the 0.2-0.8 range, as probabilities of "old" responses.

For the exemplar model, the distances between the item and all the exemplars in a category were computed. There were two possible options for classification: either summing the distance values for each category, or finding the best match in each category. Both methods were implemented and analyzed separately. In the exemplar-sum model, the sums of distance

Table 12. Mean probabilities predicted by each model for studied and unstudied items in the feature frequency and similarity experiments.

	feature frequency		Similarity	
	studied	**non-studied**	**studied**	**non-studied**
Human	0.85	0.49	0.86	0.75
CDZ-VIS	0.80	0.48	0.80	0.70
Prototype	0.56	0.61	0.80	0.69
Exemplar-sum	0.66	0.63	0.80	0.63
Exemplar-match	0.80	0.62	0.80	0.58

values were compared, and the item was classified into the category with the smallest sum. For each test item, a category and its typicality score (sum of discrete similarities) were obtained. Typicality scores ranged from 0 to 16. The typicality values were scaled to the 0.2-0.8 range, as probabilities of "old" responses.

In the exemplar-match model, the item was classified into the category with the best matching member. For each test item, a category and its typicality score (the discrete similarity of the item to the best matching member of the category) were obtained. Typicality scores ranged from 0 to 4. The typicality values were scaled to the 0.2-0.8 range, as probabilities of "old" responses.

Using these probability values, old/new decisions were made by each model. The mean responses for the studied and non-studied items were computed. Table 12 shows the predicted probabilities for each condition. A repeated-measures ANOVA was performed to test the difference between responses, with study condition (studied, non-studied) as the with-in subjects variable, and model type (human, CDZ-VIS, Prototype, Exemplar-sum, Exemplar-match) as the between subjects variable.

For both experiments, there was a significant difference between responses for studied and non-studied items (feature frequency experiment:

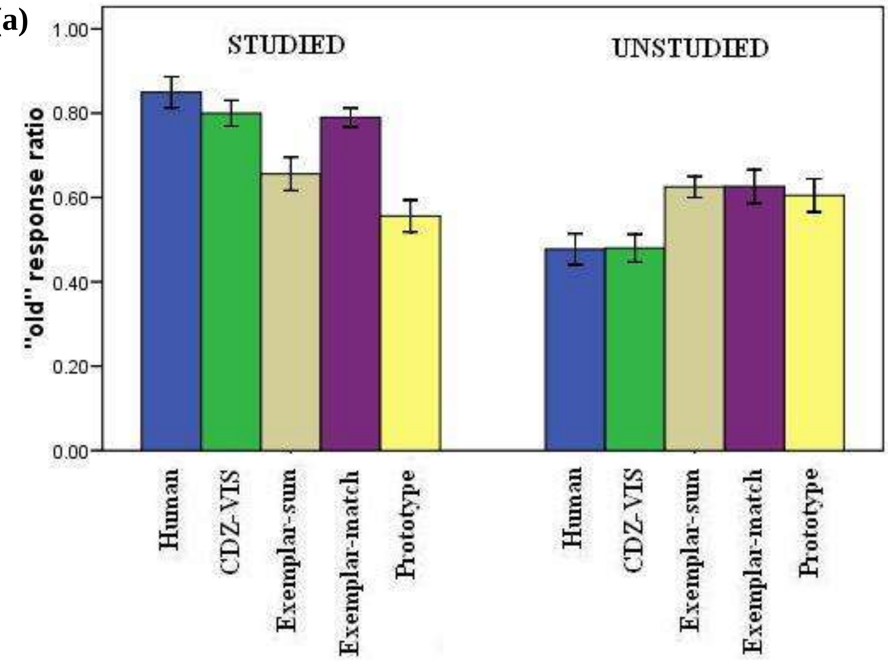
$F(1,95)=56.76$, $p<0.0001$, $\eta^2=0.37$; similarity experiment: $F(1,95)=48.71$, $p<0.0001$, $\eta^2=0.34$). There was a significant main effect of model type ($F(4,95)= 3.68$, $p<0.01$, $\eta^2=0.13$), and an interaction effect study condition*model type ($F(4,95)=13.17$, $p<0.0001$, $\eta^2=0.36$) for the feature frequency experiment, but not for the similarity experiment

As shown in Table 12 and Figure 22, exemplar-sum and prototype models predicted very low old response probabilities for studied items of the feature frequency experiment, with 0.19 and 0.29 lower probabilities than the human data, respectively. Both models compute the overall similarity of the item to all members of the category, and studied items do not necessarily result in high overall similarity values, since even a single exact match is enough to recognize a studied item. This is why exemplar-match model performed much better than these two: it finds the best match to the test item, and uses this match value to make its old/new decision.

For non-studied items, the prototype, exemplar-sum, and exemplar-match models make higher predictions than human and CDZ-VIS. As shown in Chapter 3, there was a decrease in responses with decreasing feature frequencies. An analysis of the interaction of model type and # of FRFs will be presented separately.

For the similarity experiment, the predictions of the models were exactly the same for studied items. This homogeneity in responses is a result of the homogeneity of the feature frequencies: all features were repeated exactly the same number of times. Even though the prototype and exemplar-match models computed the overall similarity, this value was high for studied items, since items were more closely gathered around the prototypes, and the match between the test items and prototypes was high. As in the feature

(a)



(b)

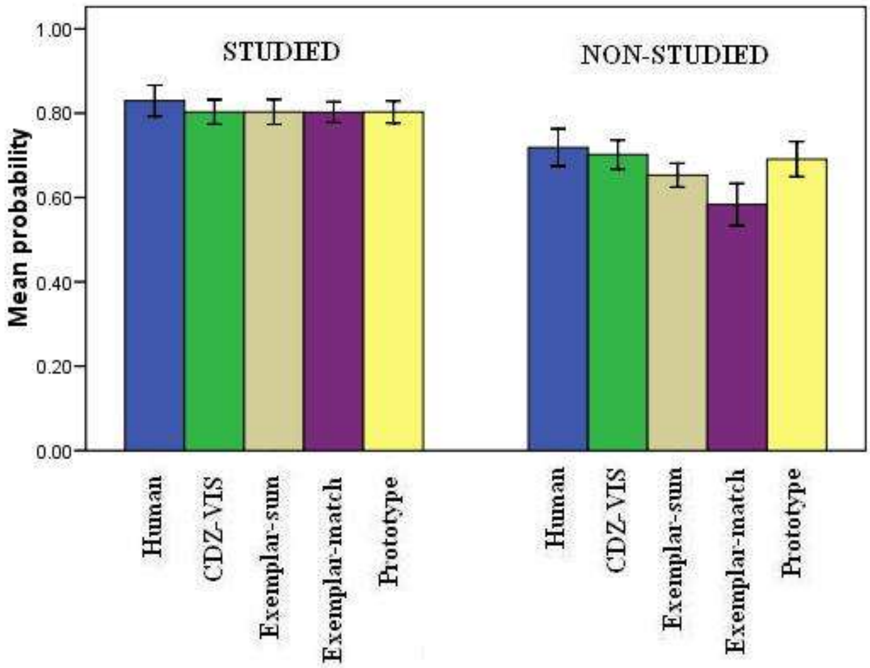


Figure 22. Comparison of model responses, according to Table 12. Error bars indicate standard error of the mean. (a) Feature frequency experiment. (b) Similarity experiment.

frequency experiment, the exemplar-match model assigns the highest probability when it finds an exact match, so it makes the same decision.

For non-studied items, the exemplar-match model was less successful than the other models. Since it considers the match between the test item and studied items, the prediction decreases as the similarity of the test item to each of the studied items decreases. However, as reported in Chapter 3, this was not the case for humans. As can be seen in Table 12 and Figure 22, prototype and exemplar-sum models performed much better. Taking the overall similarities into account increases the predicted probabilities.

Another repeated-measures ANOVA was performed to test the interaction of model type and the number of FRFs for unstudied objects in the feature frequency experiment. The effect of the number of FRFs on unstudied objects is very important for the present study: it had a very robust and strong effect on false alarms for human participants. In this analysis, the number of FRFs (0, 1, 2) was the within-subjects variable and the model type (human, CDZ-VIS, Prototype, Exemplar-sum and Exemplar-match) was the between-subjects variable. There was a main effect of the number of FRFs ($F(2, 190)= 30.13$, $p<0.0001$, $\eta^2= 0.24$), and a main effect of the model type ($F(4,95)=3.40$, $p<0.05$, $\eta^2= 0.12$). There was no interaction effect between the number of FRFs and model type. The difference between models is more related to the overall differences in absolute values. However, the more important thing is the increase in "old" response ratios as the number of FRFs on the item increases. This trend is seen in all models except the exemplar-match model (Figure 23), which finds the best match to the test item and retrieves the degree of match. Thus, it does not take the regularities in the study stimuli into account. This is the reason for the low ratio of "old" response for items with 2 FRFs. Actually, the ratios are the same for items with 1 FRF and items with 2 FRFs, since they are both items with MDS=3. The exemplar-match model considers only the similarity of the item to the items in the study phase, whereas other models

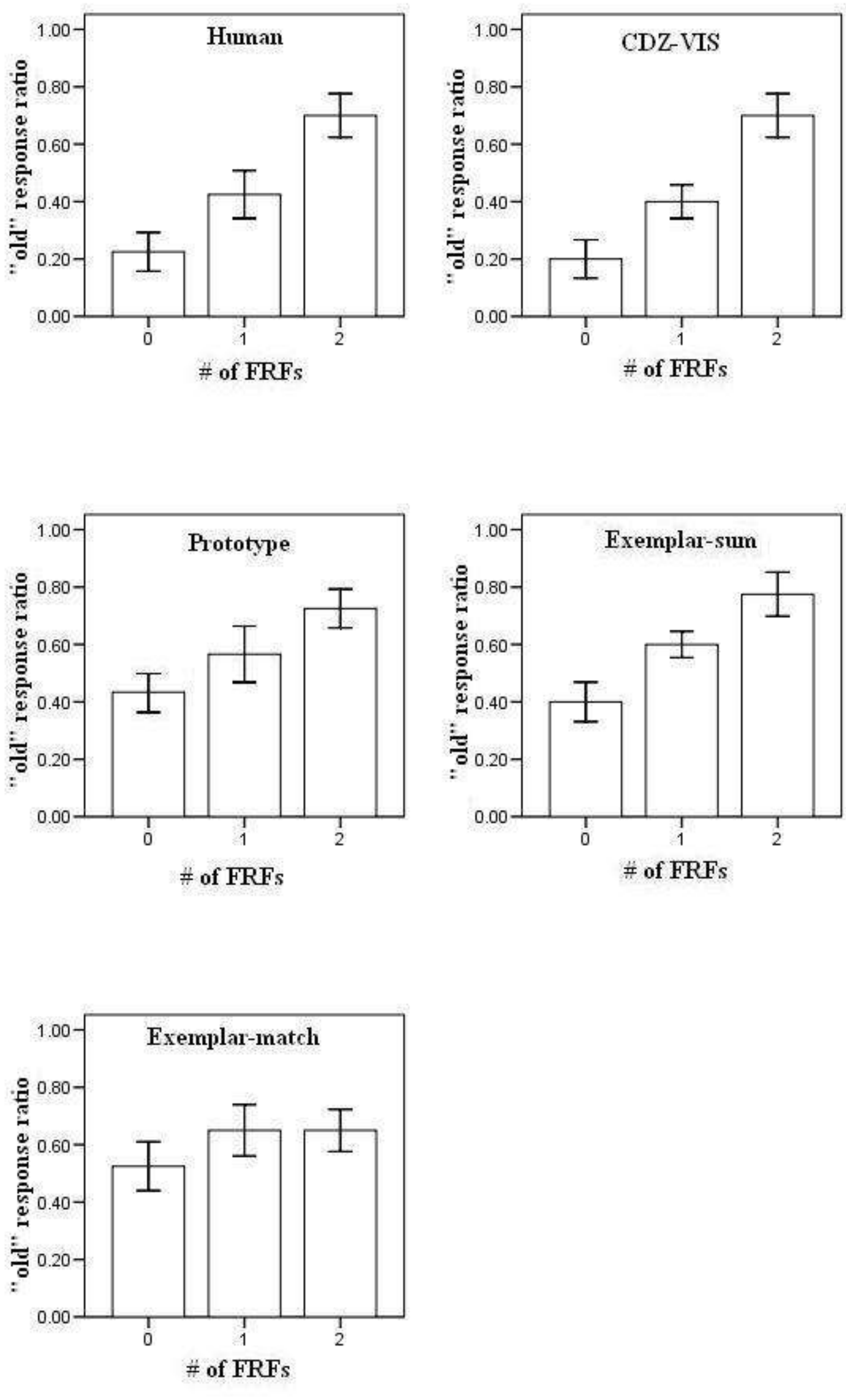


Figure 23. Ratios of old responses with respect to the number of FRFs on the unstudied test item for each model.

consider the overall regularities in the study stimuli, by computing the sum of similarities, or a prototype representing the central tendency. Thus, all the models that consider the overall regularities in the study stimuli can demonstrate the increase in the “old” responses for unstudied objects. The CDZ-VIS model displays the same trend, but with a better prediction for the items with 0 FRF. Prototype and exemplar-sum models predict a smoother increase in ratios of old responses with the number of FRFs, since they only consider the degree of match to the overall regularities in the study stimuli. CDZ-VIS predicts a sharper increase, since the low ratios of old responses are hypothesized to be a result of the matching process between the recollected features and the features of the presented test object. Recollection forces a stricter match between the recollected features and the features of the presented test object. However, recollection is not always possible and the decision is based on familiarity, as in the similarity experiment.

In general, prototype and exemplar-sum models predicted low probability of “old” responses for studied items. In other words, they made more “misses” than human participants. CDZ-VIS and exemplar-match models considered the exact match to items in the study phase, so they predicted higher probabilities. Thus, there should be a mechanism to match the test item to individual items in the study phase. One can argue that there could be a separate mechanism for object recognition and categories make contribution only for the identification of unstudied items. Individual matches with the study items do not seem to be affecting responses for unstudied items, since the exemplar-match model showed the worst performance for unstudied items. Instead of separate mechanisms, a single mechanism is suggested in the CDZ-VIS model, where the interactions between bottom-up activations from the presented test items and top-down activations from categories constructed during the study phase result in the observed pattern in human participants.

The balance between representing individual items and representing regularities in the study stimuli is a fundamental problem in categorization and recognition. In fact, proponents of the prototype models criticize exemplar models for not being able to show typicality effects of categories and proponents of the exemplar models criticize prototype models for not being able to account for effects resulting from representations of individual items. The analyses in this section showed that both models demonstrated similar performance, supporting the findings of Casale and Ashby (2008), with slight changes in absolute values, but mainly the same pattern of results. However, if the exemplar model uses a more matching-based strategy instead of using an overall summation, as suggested by Smith (2005), it shows a very different performance, which was better for identifying studied items but not compatible with human data for unstudied items.

6.3 The CDZ-VIS model and associative processes

The second type of explanations for the false memory phenomena involves the activation of existing associations during encoding, which can increase false recognition and recall during retrieval. The dominant models among the proponents of this explanation are artificial neural network models. The PDP has been one of the most influential models in many areas, including categorization and object recognition. The core principles of the PDP model will be discussed qualitatively with respect to the principles of the CDZ-VIS model.

6.3.1 The CDZ-VIS model and the PDP model

The PDP model, developed by Rumelhart and Zipser (1985) to explain semantic cognition, was later modified and analyzed by Rogers and McClelland (2004) to demonstrate various cognitive phenomena.

One issue with the Artificial Neural Network (ANN) models of object recognition and categorization is that each phenomenon is explained with a different model. For example, in order to explain memory distortions, McClelland (1995) combines two models which he calls a "Trace Synthesis Model". A similar property of the model to the CDZ-VIS model is the intra-layer inhibitory connections and inter-layer excitatory connections. Also, the activation of categories leads to divergent activations towards individual items and their properties (or features). Since every item representation under the category is activated, all of them will activate their properties at the lower layer. Intra-layer inhibition avoids maximal activation of any single representation, and causes top-down partially-activated representations at the feature layer. McClelland states that generalization is possible because of these partial activations. In the CDZ-VIS model, intra-layer inhibition is possible only if a particular representation is strong enough to inhibit others. The strength is determined by the number of representations at lower layers converging to it. An unstudied item can be recognized as "old" if it belongs to a category with strong divergent connections. An item representation will not receive enough activation if the category it belongs to is suppressed by inhibition from stronger categories in the same episodic representation.

Another similar aspect of the "Trace Synthesis Model" is the bottom-up partial activation of the item representations because of matching features between the presented stimulus and stored representations. McClelland states that this is both an advantage and disadvantage for the network. It is an advantage since unstudied items can activate existing representations so that some information can be retrieved from memory that might help identify the item. It might also create a disadvantage since the retrieved representation will not be an exact match to the presented stimuli, and the retrieved information might not be helpful, even misguiding. This property is also valid for the CDZ-VIS theory, and the false memory effects obtained in the similarity experiment are explained with this property. Thus, a mechanism for generalization can be the basis for false memories, as occurs in both models.

McClelland also attributes false memories to recollection problems. Since representations of items are just connections of lower-layer properties, each connection should be activated if a specific property is to be recollected. If these connections are not strong enough, these details cannot be recollected, and false memory effects may occur. In the CDZ-VIS model, the recognition decision depends on the partial activations in the object layer, if the features converging to them cannot be retrieved. There is evidence in the neuroscience literature that old people rely on familiarity when specific details of the item cannot be recollected (Plancher, Guyard, Nicolas, & Piolino, 2009).

McClelland (1995) designed an experiment to produce data for his model, in which they presented sentences and then asked participants to complete fragments of sentences. In order to prove that the Trace Synthesis Model can simulate data from this experiment, they designed a network with parameters to adjust the network to the data and with additional assumptions to the original model like randomly eliminating some connections and adding random noise to input units. This modified model simulated the data successfully. Then, he introduced Hinton’s model for gradual learning, and combined the two models, obtaining a dual memory system. The CDZ-VIS theory involves similar mechanisms as in the Trace Synthesis Model, and also a categorization mechanism, however, without combining several models together and adjusting parameters for each phenomenon. It has a basic mechanism, whose behavior changes depending on the presented stimuli. It can explain various phenomena in the same structure, without building a hybrid from existing models. In this regard, the Trace Synthesis Model can be considered as an instance of the CDZ-VIS theory where items cause partial activations in the item representations. How regularities in the stimulus set can affect these partial activations remains unexplained in McClelland's Trace Synthesis Model of recognition.

Rogers and McClelland (2004) propose a modified version of Rumelhart's PDP model as a model of semantic cognition and provides detailed analysis of its behavior. There are six core principles of their model:

- Predictive Error-Driven Learning
- Sensitivity to Coherent Covariation
- Similarity-Based Distributed Representation
- Convergence of Influences on Representations and Connection Weights
- Gradual, Structure-Sensitive Learning
- Activation-Based Representation of Novel Objects

In predictive error-driven learning, the network calculates an output for the given input, and compares the predicted output to the actual output. In such a learning mechanism, the network needs several examples to learn the input-output mappings. The training takes such a long time that a semantic learning task can take as much as 18.000 epochs where the whole set of input-output pairs is studied 18.000 times. The network tries to approximate a function which maps presented inputs to outputs. First of all, the network needs many examples to make a good approximation. Second, the network needs feedback, in terms of the actual output. Third, it needs to be trained thousands of times to learn a simple set of input-output mappings. In the CDZ-VIS theory, learning takes place when synchronously firing neurons build connections to ("convergence to", in Damasio's (1989) terminology) upper layer associative neurons. Such a network does not need an input-output pairing: the convergence of the network results in representations of the regularities in the input. In other words, the network is self-organizing. External feedback is not required. In terms of the computation time, CDZ-VIS is much superior, since a single presentation of the item is enough for its representation to be formed. This is more compatible with learning in the brain, which does not require the presentation of an item several times to be recognized. Multiple presentations make connections stronger, but a single presentation is enough to influence later recall and recognition.

The coherent covariation principle describes the sensitivity of the network to covariations in the input. CDZ-VIS is compatible with this principle since the covariations are reflected in the similarity measure. Co-occurring features cause item representations to form closer and converge to upper layer category CDZs. In the PDP model, covariations are reflected in the connection weights, and the regularities are represented in a distributed manner, not in CDZ-like localized units. However, recent evidence from neuroscience studies shows that there are local regions that are sensitive to co-varying features in the input (Hommel, 2004).

Actually, the main difference between the PDP approach and CDZ-VIS theory is the distributed representations in the PDP network. In the PDP approach, items are not stored locally but the information (feature bindings, regularities) is stored in the connections. In CDZ-VIS, items have their corresponding CDZs, which are formed as a result of convergence of lower layer representations. This structure is more compatible with recent evidence from neuroscience, which shows that there are local regions sensitive to particular stimuli, and there is a hierarchical structure from low-level feature representations to more abstract representations of objects and categories, each of them being local representations (Rolls & Deco, 2002). A representation being local does not mean that it does not respond to variations of the stimuli. If the features of the presented item activate the features of an existing representation, this previously built representation can be partially activated. It also means that a network with localized representations can demonstrate partial activations and spread of activation. Thus, similar principles operate in CDZ-VIS and ANNs, but with different computations, and CDZ-VIS is suggested to be a more biologically-plausible explanation of the behavioral phenomena than the PDP approach. In fact Rogers and McClelland also mention that the PDP is only one of the possible mechanisms that can explain existing phenomena.

The reason for employing the distributed structure as a basis of their models as stated by Rogers and McClelland is that it enables efficient representations of items with a limited number of units. Besides, learning of an item affects all other representations to the degree of their similarity. However, this property is not unique to distributed representations. In the CDZ-VIS model, representations of similar items are closer to each other, and when an unstudied item is presented to the network, since existing representations are activated to the degree of their similarities, their connections are also updated. Thus, one does not need to have a distributed representation and error-driven learning to have predictive power and similarity-based activations.

The "Convergence of Influences on Representations and Connection Weights" principle is related with the hidden units. These units are crucial for the backpropagation of the error, and thus for the learning in the network. Any regularity in the input can be detected and represented by the hidden units. In the CDZ-VIS theory, CDZs behave like hidden units. Regularities in a stimulus set lead to synchronous activation of item representations, which in turn leads to convergence of these representations to CDZs. There can be many layers of CDZs, where layer location is determined by the similarity of the converging items. However, similarity is not required for synchronous activation: any region synchronously active can converge to upper layer CDZs. Thus, the mechanisms for representing regularities are much more flexible in the CDZ-VIS model. The number of hidden layers in a PDP can increase computational load in the network significantly, and the required number of hidden units depends on several factors, including number of training cases, number of input units, number of output units, the training algorithm, etc. In CDZ-VIS, the CDZs are not hidden units, though, since every CDZ is a representation itself. The number of CDZs is determined by the number of convergences of item representations to upper-layer CDZs. Thus, there is no need for a pre-defined number of CDZ.

There is also a “convergence principle” defined by Rogers and McClelland. It corresponds to the connection of several units to a single unit, where representations are patterns of activations throughout the network. In the CDZ-VIS theory, convergence has a different meaning, where every converging representation connects to its own representing CDZ. Thus, there is one CDZ for every converging item representation, opposite to the idea in the PDP model where all the input and hidden units converge to a single output representation at some point in the network. One thing in common is the many-to-one relation in general.

Gradual, structure-sensitive learning refers to the slight modifications of the connections with each input to the network. In the CDZ-VIS theory, each stimulus presentation causes formation of an object-representing CDZ, and category-representing CDZs if common features are activated with the existing representations. Whether this change in the network is a small or big change depends on the definition of a small change. Formation of a CDZ means new connections develop among different regions of the brain. The amount of change depends on the number of common features between the presented stimulus and existing representations. If there are many common features, many CDZs will form, which can be considered as a big change. If no common features exist, there will be only a single CDZ formation for the item itself or just connections from the features of the presented stimuli to an existing representation if the object was encountered before. Thus, the amount of change in the network depends on the number of activations caused by the presented input. In the PDP network, similar processes take place like the bigger change in connection weights for more activated units, but the computations are very different. Thus, the CDZ-VIS theory adheres again to the same principle, however, with a different network structure which is suggested to be biologically more plausible.

The last principle is the activation-based representation of novel items. When a novel item is introduced to the network, it activates a specific output unit in the network. The representation of the item is the pattern of activations in the network. Since the learning in the network depends on the slow weight updates each time, the introduction of a new item might cause an undesired large change in the weights. Thus, Rogers and McClelland suggests a dual-processing system with fast-learning and slow-learning components. The fast-learning component quickly captures the input-output pair obtained from the environment, and trains the slow-learning network in time for a more stable learning experience. This is called “complementary learning systems” by Rogers and McClelland. Even though the fast-learning component learns the pair first, the input is first introduced to the slow-learning network to retrieve the pattern of activation created by the input. Thus, the input is first evaluated by prior experience, but the learning from the new experience occurs in a two-stage process.

In the CDZ-VIS theory, learning takes place when synchronously firing neurons converge to associative neurons, called CDZs. The stability problem in the artificial neural networks is not applicable in such a learning mechanism, since the representations are not distributed in the sense that ANNs employ. Thus, there is no need for a dual processing system. Convergence to a new CDZ does not cause large changes in the whole network every time a new stimulus is introduced. Rather, this new CDZ can cause activations, depending on the similarity of the existing representations in the neighborhood, and the activations caused by the spread of activation to the previously built connections. Thus, only connections to the related representations are updated. Unnecessary update of weights is not possible by the nature of the mechanism. Still, the network can display almost all the principles supported by ANNs which are considered essential to explain psychological and neuroscience findings.

6.4 Summary

In this chapter, the CDZ-VIS model was qualitatively and quantitatively compared to some fundamental findings and models in the false memory literature, which also involves models of categorization and recognition. When compared with the models of categorization, the CDZ-VIS model provided the closest predictions to human data. The reason for this was that the prototype and exemplar models considered only the overall similarity of the test items to the members of the categories that form during the study phase. They did not take representation of the individual items into account, except the exemplar-match model. Thus, to make better predictions, a model of object representations should consider both individual item representations and category representations, or regularities. The CDZ-VIS model builds a hierarchical representation of items in the study phase from low-level features to episodic representations, which involves both item and category representations, as well as feature and episodic representations. During the test phase, these representations interact with each other, using both bottom-up (convergent) and top-down (divergent) connections, and the resulting activations in the object layer determine the result of the old/new decision. This mechanism is compatible with recent neuroscience findings, from hierarchical representations to re-activation of previously built representations during recognition (Danker & Anderson, 2010). The CDZ-VIS model was also compared to the PDP model as discussed in Rogers and

McClelland (2004). While the underlying principles were compatible, the computations were fundamentally different in two models: representations in CDZ-VIS are local whereas representations in PDP are distributed. This difference changes the whole processing framework.

CHAPTER 7

DISCUSSION AND CONCLUSION

In this chapter, the results of the experiments and modeling study are summarized, and the implications of the proposed model for the literature on visual object representations are discussed. The chapter ends with limitations of the study and suggestions for future studies.

7.1 Summary of the study

This study investigated the formation and structure of visual object representations. The study consisted of two parts: a behavioral experiment and a modeling study. The aim of the first study was to test specific hypotheses about possible factors affecting the formation and retrieval of visual object representation. In the second part, a model of the formation of visual object representations was proposed based on the findings from neuroscience and cognitive psychology literature, and a validation study was performed in Chapter 6.

The following hypothesis about the formation of visual object representations guided both behavioral and modeling studies: The repetition frequencies of object features and constraints of the neural structures result in a hierarchical representation from features to categories. The behavioral experiment tested the effect of feature repetition frequency on recognition of objects as “old” and “new”. Consequently, the modeling study investigated whether a mechanism with biologically plausible components can explain phenomena regarding perception and memory of objects. In the next section, the findings of these studies are discussed.

7.2 Behavioral experiments

Two experiments were performed to test the hypothesis that the repetition frequency and similarity of object features affects old/new recognition. The study/test, old/new recognition task was used as the experimental paradigm. A study/test old/new recognition task consists of two phases: study and test. During the study phase, a set of objects is presented to participants one-by-one. In the test phase, another set of objects is presented, and the participant is required to tell whether the object is studied or unstudied.

7.2.1 Feature frequency experiment

In the first experiment, the repetition frequency was controlled for individual features and feature combinations. As a result, some features and feature combinations were presented more frequently than other features and feature combinations during the study phase. During the test phase, old/new responses of participants were collected for each test object. The data from the experiment were analyzed with repeated-measures ANOVA. There was a main effect of study condition (studied, unstudied) and a main effect of the number of FRFs on the object. Thus, the hypothesis that the repetition frequency of features affects old/new responses was confirmed since there was a significant effect of number of FRFs on the old/new responses. There was also an interaction effect between the study condition and the number of FRFs. This interaction effect indicated that the effect of the number of FRFs was different for studied and unstudied objects. There was a significant effect of the number of FRFs on unstudied objects, but not on studied objects. For unstudied objects, as the number of FRFs on the object increased, the percentage of “old” responses for that object increased.

If a particular feature and feature combination are repeated more frequently than others during the study phase, unstudied objects in the test phase which have these features were recognized as “old”. This is a false memory phenomenon. Objects which were not presented during the study phase were remembered to be seen, which is called a false alarm in the signal detection literature.

7.2.2 Similarity experiment

In the second experiment, the effect of similarity on old/new responses was investigated. This time, the frequency of all features and feature combinations in the study phase were equal. A similarity measure was defined as MDS, which is the maximum similarity of an object with objects in the study phase. Test objects were selected according to their MDS values. Old/new responses of the participants were collected, and analyzed by repeated-measures ANOVA. Percentage of “old” responses was high in general and the difference between levels of similarity was only significant for MDS_3 and MDS_4. This difference is actually the difference between studied and unstudied objects. Thus, the similarity of the test items to the study items did not affect old/new responses. Participants found most test objects “old”, and the decrease in similarity did not decrease “old” responses.

7.2.3 Discussion of the results of behavioral experiments

The results of the experiments showed that the repetition frequency of features and feature combinations affects old/new responses. Even though the increase in the number of FRFs on the object seemed to be increasing the percentage of "old" responses, results of the similarity experiment showed that the percentage of "old" responses was high in general when the frequency of features and feature combinations are equal. Thus, the variations in feature frequency were actually "decreasing" the percentage of old responses for objects without FRFs.

According to the results of the two experiments, percentage of "old" responses for studied items was high, regardless of the similarity and feature frequency. This is the main function of memory: previously seen items should be recognized as old. On the other hand, a different pattern of responses was observed for unstudied objects.

First, when similar items are presented during the study phase, participants recognized most of the test objects as "old". What can be the benefit of recognizing unstudied objects as "old when they are similar to the previously seen objects? Even though it can be seen as an error, there are many advantages to such recognition. The most important one is that similar objects have usually similar affordances. For example, all chairs are similar to each other. If an unstudied item is similar to a chair, then it is possible that the object is sittable. Even though it may not be sittable, it is an advantage to be aware of this possibility to investigate further. Besides, if the item is similar to a dangerous entity previously encountered, it is best to be alert about it before identifying it in detail. Thus, the feeling of "oldness" is like a first step before taking the relevant action.

Second, when the feature frequencies were not equal during the study phase, the percentage of "old" responses decreased for objects without FRFs whereas it still remained high for objects with FRFs. Now, why would the recognition system prefer to call an item "new" even though it is similar to the previously seen objects, considering the discussion above? In order to answer this question, another question should be answered: what happens when the feature frequencies are equal and when the feature frequencies are different? In the similarity experiment, all objects have exactly two overlapping features with two other objects. In this case, three categories emerge from the overlapping features. These categories are equal in strength, i.e. they all contain exactly three objects. On the other hand, in the feature frequency experiment, frequencies of particular features and feature combinations were higher. Thus, many categories with different strengths emerged. For example, there exists a category resulting from *five* objects with DS=2 whereas another category resulting from *two* objects with DS=2.

The difference between the feature frequency and the similarity experiment was in terms of the study stimuli. One consequence of that difference is the number of categories which emerge as a result of the common features between the objects. In both experiments, the test objects had varying degrees of similarity to the study objects. Thus, the difference in repetition frequencies must be affecting the representations formed during the study phase. This indicates that the two experiments differed in terms of the representational structure.

Returning back to the previous question about the benefit of recognizing an object as new even though it is similar to the previously seen objects, the representational structure in the unequal feature frequencies situation should be investigated. In a context where some features are more frequent and others less frequent, why would the high frequency ones should make unstudied objects "old" even though they are unstudied? If a feature is repeating on the objects, and other features are changing, then these repeating features becomes the invariant features of these objects and the others are varying features. In natural environments, this is usually the case when the members of a category are presented. For example, in a box of apples, there can be green, red and yellow apples, small and big apples, dotted or smooth apples. However, some features are always the same, like shape and texture. These common features make it possible to recognize an apple.

Thus, in a situation where some features are more frequent and others not, the categories of objects can be learnt. The frequent features are the invariant properties of the category and the less frequent features are irrelevant to the category. Even though an unstudied object is similar to the objects in the study phase, there is no need to recognize it as “old”, since the categories are learnt, and the unstudied object does not belong to these categories. The information about the categories is important again for identifying the object’s possible functionalities or affordances. The similarity of the objects to the study objects is not enough now to infer its functionalities or affordances.

7.3 The CDZ-VIS model

The behavioral experiments imposed new constraints on a model of the structure of visual object representations in memory. It was hypothesized that a model with the principles stated in Chapter 5 can satisfy these constraints.

Evidence from the neuroscience literature provides many insights into the problem of the formation of visual object representations. First, the neural populations have competitive interactions where each neural population competes for activating as many neighbors as possible and inhibiting as many competitor populations as possible. When some categories are stronger, they have the chance to win against weaker categories. However, if the category strengths are equal, they compete with each other but no one can win.

Thus, in Chapter 4, a number of models which have competitive units were examined, and tested with data from behavioral experiments. Since these models are very specific to a single phenomenon (viewpoint and illumination invariance), they could not predict the effects observed in the behavioral experiments.

The second question is how objects and categories are organized so that categories can emerge from objects. Again, the solution might be found in the neuroscience literature. The currently accepted view on the structure of visual object representations is a hierarchical network starting from occipital region and reaching towards inferior temporal lobe, from concrete object features to abstract representations of categories. The connections in the hierarchy form by convergence from lower layer neurons to upper layer associative neurons. Thus, according to this literature, feature representations converge to object representations, and object representations converge to category representations. In addition to convergence, there is also divergence from upper layers to lower layers, which corresponds to the top-down effects of memory representations on sensory processes.

The third question is how the resulting representations are used for perceptual and memory processes. Current findings in the domain of cognitive neuroscience indicate that representations are activated bottom-up by sensory stimuli and top-down by high-level contextual information, i.e. information about the current context: task at hand, place, time, etc. Top-down effects in perception is a well-studied subject in the perception literature. Instead of purely bottom-up processing, the currently accepted view is an interaction between bottom-up and top-down processes in perception and memory.

The prominence of the Convergence-Divergence Zone Framework (Meyer and Damasio (2009) in this study among other memory frameworks in the literature was a result of its compatibility with the recent neuroscience findings explained above. The CDZ-VIS model was developed on this framework with detailed formal specifications of the encoding, retrieval and recognition processes, which were presented in Chapter 5. Since the specifications have been determined to be compatible with the most recent neuroscience findings, the level of abstraction is much lower than current explanations of recognition and categorization. If the level of abstraction were high, the variety of the phenomena which could be explained by the model would be less, since the building blocks of the model are abstractions of specific factors directly related to a specific phenomenon. When the components of the model are independent of a specific phenomenon, like models of neurons, they are not restricted to a particular phenomenon. Thus, the current level of abstraction in the CDZ-VIS model is a suitable one: it is based on the dynamics of neural populations, and it can explain behavioral phenomena.

When compared with the predictions of the prototype and exemplar models in Chapter 6, the CDZ-VIS model provided the closest predictions to human data. The prototype and exemplar models considered the overall similarity of the test items to the members of the categories that form during the study phase. The CDZ-VIS model made better predictions by taking the representations of individual objects into account, as well as the representations of regularities in the study stimuli, building a hierarchical representation from low-level features to episodic representations, involving feature, object, category, and episodic representations.

7.4 Conclusion

In this dissertation, behavioral experiments to test the effects of feature frequency and similarity on old/new recognition were performed and a model of the formation and structure of visual object representations has been developed. The design of the CDZ-VIS model is based on principles derived from current findings from cognitive neuroscience, and the results of the behavioral experiments provided new constraints for the model to satisfy. Because of the fact that the model has been built on neural dynamics, such as competition, inhibition and spread of activation, it provided a convenient level of explanation for the formation of object representations and recognition of objects. There have been various studies on the formation of visual object representations under different disciplines of cognitive science, but there has been little effort for building a mechanism which integrates current findings from these disciplines. The proposed model is a step towards making this integration and providing a common ground for producing and testing new hypotheses about object representations.

7.5 Limitations and suggestions for further studies

The results of the behavioral experiments and the CDZ-VIS model of visual object representations bring new perspectives to the formation of categories from features and object in a hierarchical system. The CDZ-VIS model is still under development, in terms of the contextual influences, and the interaction of the visual object representations with representations in other modalities, involving auditory, tactile and action features. The hypothesis that every feature in the scene is encoded in the same way like visual features has already been tested (Eren-Kanat & Hohenberger, 2011). Preliminary results from these studies indicated that old/new responses are associated only with visual object features. Other features in the scene, like spatial position and action effects, are relevant to other processes, like whether the object is displayed previously in the same position or not, or whether the object displayed the same action effect. Thus, it might be possible that representations in different modalities are formed in the same way, but activated by different task relevant contextual features. For example, visual object representations might be activated in contexts where the identity of the object is relevant, but not when the position of the object is relevant.

REFERENCES

Ariga, A., Yokosawa, K., & Ogawa, H. (2007). Object-based attentional selection and awareness of objects. *Visual Cognition, 15*(6), 685-709. doi: 10.1080/13506280601016967

Azimian-Faridani, N., & Wilding, E. L. (2004). An event-related potential study of the revelation effect. *Psychonomic Bulletin & Review, 11*(5), 926-931.

Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*(4), 629.

Barsalou, L. W. (2003). Abstraction in perceptual symbol systems. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 358*(1435), 1177-1187. doi: 10.1098/rstb.2003.1319

Beck, D. M., & Kastner, S. (2009). Top-down and bottom-up mechanisms in biasing competition in the human brain. *Vision Research, 49*(10), 1154-1165.

Biederman, I. (1987). Recognition by components: A theory of human image understanding. *Psychological Review, 94*(2), 115-147.

Brainerd, C. J., & Kingma, J. (1984). Do children have to remember to reason? A fuzzy-trace theory of transitivity development. *Developmental Review, 4*(4), 311-377.

Brainerd, C. J., & Wright, R. (2005). Forward association, backward association, and the false-memory illusion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(3), 554.

Bremner, A. J., Bryant, P., Mareschal, D., & Volein, A. (2007). Recognition of complex object-centred spatial configurations in early infancy. *Visual Cognition, 15*(8), 896-926. doi: 10.1080/13506280601029739

Bussey, T. J., & Saksida, L. M. (2005). Object memory and perception in the medial temporal lobe: an alternative approach. *Current Opinion in Neurobiology, 15*(6), 730-737. doi: 10.1016/j.conb.2005.10.014

Casale, M. B., & Ashby, F. G. (2008). A role for the perceptual representation memory system in category learning. *Attention, Perception, & Psychophysics, 70*(6), 983-999.

Colzato, L. S., Raffone, A., & Hommel, B. (2006). What do we learn from binding features? Evidence for multilevel feature integration. *Journal of Experimental Psychology-Human Perception and Performance, 32*(3), 705-716. doi: 10.1037/0096-1523.32.3.705

Cutzu, F., & Edelman, S. (1996). Faithful representation of similarities among three-dimensional shapes in human vision. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(21), 12046-12050.

Damasio, A. R. (1989). The brain binds entities and events by multiregional activation from convergence zones. *Neural Computation, 1*(1), 123-132.

Danker, J. F., & Anderson, J. R. (2010). The Ghosts of Brain States Past: Remembering Reactivates the Brain Regions Engaged During Encoding. *Psychological Bulletin, 136*(1), 87-102. doi: 10.1037/a0017937

Daprati, E., & Sirigu, A. (2006). How we interact with objects: learning from brain lesions. *Trends in cognitive sciences, 10*(6), 265-270.

Davachi, L. (2006). Item, context and relational episodic encoding in humans. *Current Opinion in Neurobiology, 16*(6), 693-700. doi: 10.1016/j.conb.2006.10.012

Dawson, M. R. W. (2004). *Minds and machines: Connectionism and psychological modeling*: Wiley-Blackwell.

Dijkerman, H. C., Milner, A. D., & Carey, D. P. (1996). The perception and prehension of objects oriented in the depth plane .1. Effects of visual form agnosia. *Experimental Brain Research, 112*(3), 442-451.

Duzel, E., Habib, R., Rotte, M., Guderian, S., Tulving, E., & Heinze, H. J. (2003). Human hippocampal and parahippocampal activity during visual associative recognition memory for spatial and nonspatial stimulus configurations. *Journal of Neuroscience*, *23*(28), 9439-9444.

Edelman, S., Cutzu, F., & DuvdevaniBar, S. (1996). *Similarity to reference shapes as a basis for shape representation*.

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience, 30*, 123-152. doi: 10.1146/annurev.neuro.30.051606.094328

Epstein, R., Graham, K. S., & Downing, P. E. (2003). Viewpoint-specific scene representations in human parahippocampal cortex. *Neuron, 37*(5), 865-876.

Fahle, M., & Poggio, T. (2002). *Perceptual learning*: MIT Press.

Fiete, I. R., Senn, W., Wang, C. Z. H., & Hahnloser, R. H. R. (2010). Spike-time-dependent plasticity and heterosynaptic competition organize networks to produce long scale-free sequences of neural activity. *Neuron, 65*(4), 563-576.

Földiák, P. (1991). Learning invariance from transformation sequences. *Neural Computation, 3*(2), 194-200.

Gerstner, W. (2000). Population dynamics of spiking neurons: fast transients, asynchronous states, and locking. *Neural computation, 12*(1), 43-89.

Glanzer, M., Adams, J. K., Iverson, G. J., & Kim, K. (1993). The regularities of recognition memory. *Psychological Review, 100*(3), 546.

Hanna, A., & Remington, R. (1996). The representation of color and form in long-term memory. *Memory & Cognition, 24*(3), 322-330.

Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences, 8*(11), 494-500.

Hommel, B., & Colzato, L. S. (2009). When an object is more than a binding of its features: Evidence for two mechanisms of visual feature integration. *Visual Cognition, 17*(1-2), 120-140. doi: 10.1080/13506280802349787

Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review, 99*(3), 480-517.

Johnson, M. K. (1988). Discriminating the origin of information.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*(2), 175-219.

Keizer, A. W., Colzato, L. S., & Hommel, B. (2008). Integrating faces, houses, motion, and action: Spontaneous binding across ventral and dorsal processing streams. *Acta Psychologica, 127*(1), 177-185. doi: 10.1016/j.actpsy.2007.04.003

Kelley, C. M., & Jacoby, L. L. (1990). The construction of subjective experience: Memory attributions. *Mind & Language, 5*(1), 49-68.

Kohler, S., Black, S. E., Sinden, M., Szekely, C., Kidron, D., Parker, J. L., Bronskill, M. J. (1998). Memory impairments associated with hippocampal versus parahippocampal-gyrus atrophy: an MR volumetry study in Alzheimer's disease. *Neuropsychologia, 36*(9), 901-914.

Kosslyn, S. M. (2005). Mental images and the brain. *Cognitive Neuropsychology, 22*(3-4), 333-347. doi: 10.1080/02643290442000130

Kroll, N. E. A., Knight, R. T., Metcalfe, J., Wolf, E. S., & Tulving, E. (1996). Cohesion failure as a source of memory illusions. *Journal of Memory and Language, 35*, 176-196.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*: University of Chicago press.

Landau, J. D. (2001). Altering the balance of recollection and familiarity influences the revelation effect. *American Journal of Psychology, 114*(3), 425-437.

Lee, H., & Vecera, S. P. (2005). Visual cognition influences early vision - The role of visual short-term memory in amodal completion. *Psychological Science, 16*(10), 763-768.

Li, J. L., & Yeh, S. L. (2007). New objects do not capture attention without a top-down setting: Evidence from an inattentional blindness task. *Visual Cognition, 15*(6), 661-684. doi: 10.1080/13506280600926695

Luck, S. J., Girelli, M., McDermott, M. T., & Ford, M. A. (1997). Bridging the gap between monkey neurophysiology and human perception: An ambiguity resolution theory of visual selective attention. *Cognitive Psychology, 33*(1), 64-87.

Marr, & David. (1982). *Vision : a computational investigation into the human representation and processing of visual information*. San Francisco ; New York: W.H. Freeman.

Matsumoto, N., Okada, M., Sugase-Miyamoto, Y., Yamane, S., & Kawano, K. (2005). Population dynamics of face-responsive neurons in the inferior temporal cortex. *Cerebral Cortex, 15*(8), 1103.

McClelland, J. L. (1995). Constructive memory and memory distortions: a parallel-distributed processing approach. *Memory distortion: How minds, brains, and societies reconstruct the past*, 69-90.

McClelland, J. L. (2009). The place of modeling in cognitive science. *Topics in Cognitive Science, 1*(1), 11-38.

Meyer, K., & Damasio, A. (2009). Convergence and divergence in a neural architecture for recognition and memory. *Trends in Neurosciences, 32*(7), 376-382. doi: 10.1016/j.tins.2009.04.002

Mishkin, M., & Ungerleider, L. G. (1982). Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behavioural Brain Research, 6*(1), 57-77.

Mulligan, N. W., & Lozito, J. P. (2006). An asymmetry between memory encoding and retrieval - Revelation, generation, and transfer-appropriate processing. *Psychological Science, 17*(1), 7-11.

Nosofsky, R. M., & Johansen, M. K. (2000). Exemplar-based accounts of" multiple-system" phenomena in perceptual categorization. *Psychonomic Bulletin and Review, 7*(3), 375-402.

Nosofsky, R. M., Little, D. R., Donkin, C., & Fific, M. (2011). Short-Term Memory Scanning Viewed as Exemplar-Based Categorization. *Psychological Review, 118*(2), 280-315. doi: 10.1037/a0022494

Nosofsky, R. M., & Stanton, R. D. (2005). Speeded classification in a probabilistic category structure: contrasting exemplar-retrieval, decision-boundary, and prototype models. *Journal of Experimental Psychology: Human Perception and Performance, 31*(3), 608.

Plancher, G., Guyard, A., Nicolas, S., & Piolino, P. (2009). Mechanisms underlying the production of false memories for famous people's names in aging and Alzheimer's disease. *Neuropsychologia, 47*(12), 2527-2536.

Ranganath, C., & D'Esposito, M. (2005). Directing the mind's eye: prefrontal, inferior and medial temporal mechanisms for visual working memory. *Current Opinion in Neurobiology, 15*(2), 175-182. doi: 10.1016/j.conb.2005.03.017

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience, 2*(11), 1019-1025.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(4), 803.

Roediger, H. L., Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review, 8*(3), 385-407.

Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*: The MIT Press.

Rogers, T. T., & McClelland, J. L. (2005). A parallel distributed processing approach to semantic cognition: Applications to conceptual development. In L. Gershkoff-Stowe & D. H. Rakison (Eds.), *Building Object Categories In Developmental Time* (pp. 335-387). Psychology Press.

Rolls, E. T., & Deco, G. (2002). *Computational neuroscience of vision*. Oxford University Press.

Rolls, E. T., & Milward, T. (2000). A model of invariant object recognition in the visual system: Learning rules, activation functions, lateral inhibition, and information-based performance measures. *Neural Computation, 12*(11), 2547-2572.

Rolls, E. T., & Stringer, S. M. (2006). Invariant visual object recognition: A model, with lighting invariance. *Journal of Physiology-Paris, 100*(1-3), 43-62. doi: 10.1016/j.jphysparis.2006.09.004

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of experimental psychology: General, 104*(3), 192.

Rosch, E. (1999). Principles of categorization. In E. Margolis & S. Laurence (Eds.), *Concepts: core readings*, (pp. 189–206). MIT Press.

Rumelhart, D. E., & McClelland, J. L. (1988). *Parallel distributed processing*. MIT Press.

Rumelhart, D. E., & Zipser, D. (1985). Feature discovery by competitive learning. *Cognitive Science, 9*(1), 75-112.

Schmidt, S. R. (1996). Category typicality effects in episodic memory: Testing models of distinctiveness. *Memory & cognition, 24*(5), 595-607.

Shinkareva, S. V., Mason, R. A., Malave, V. L., Wang, W., Mitchell, T. M., & Just, M. A. (2008). Using fMRI Brain Activation to Identify Cognitive States Associated with Perception of Tools and Dwellings. *Plos One, 3*(1). doi: e139410.1371/journal.pone.0001394

Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends in cognitive sciences, 12*(5), 182-186.

Slotnick, S. D., & Schacter, D. L. (2004). A sensory signature that distinguishes true from false memories. *Nature Neuroscience, 7*(6), 664-672. doi: 10.1038/nn1252

Smith, J. D. (2005). Wanted: a new psychology of exemplars. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expé rimentale, 59*(1), 47.

Smith, R. E., & Hunt, R. R. (1998). Presentation modality affects false memory. *Psychonomic Bulletin & Review, 5*(4), 710-715.

Smith, S. M., Ward, T. B., Tindell, D. R., Sifonis, C. M., & Wilkenfeld, M. J. (2000). Category structure and created memories. *Memory & Cognition, 28*(3), 386-395.

Summerfield, J. J., Lepsien, J., Gitelman, D. R., Mesulam, M. M., & Nobre, A. C. (2006). Orienting attention based on long-term memory experience. *Neuron, 49*(6), 905-916. doi: 10.1016/j.neuron.2006.01.021

Thompson-Schill, S. L. (2003). Neuroimaging studies of semantic memory: inferring. *Neuropsychologia, 41*(3), 280-292.

Tyler, C. W., Likova, L., & Nicholas, S. (2009). Decoding neural response dynamics in the human brain. *Perception, 38*, 107-108.

Underwood, B. J. (1965). False recognition produced by implicit verbal responses. *Journal of Experimental Psychology, 70*(1), 122.

Usher, M., Schuster, H. G., & Niebur, E. (1993). Dynamics of populations of integrate-and-fire neurons, partial synchronization and memory. *Neural Computation, 5*(4), 570-586.

Verde, M. F., & Rotello, C. M. (2003). Does familiarity change in the revelation effect? *Journal of Experimental Psychology-Learning Memory and Cognition, 29*(5), 739-746. doi: 10.1037/0278-7393.29.5.739

Verhoef, B. E., Kayaert, G., Franko, E., Vangeneugden, J., & Vogels, R. (2008). Stimulus similarity-contingent neural adaptation can be time and cortical area dependent. *The Journal of Neuroscience, 28*(42), 10631-10640.

Wallis, G., & Rolls, E. T. (1997). Invariant face and object recognition in the visual system. *Progress in Neurobiology, 51*(2), 167-194.

Watkins, M. J., & Peynircioglu, Z. F. (1990). The revelation effect: When disguising test items induces recognition. *Journal of Experimental Psychology-Learning Memory and Cognition, 16*(6), 1012-1020.

Wheeler, M. E., & Buckner, R. L. (2003). Functional dissociation among components of remembering: Control, perceived oldness, and content. *Journal of Neuroscience, 23*(9), 3869-3880.

Wixted, J. T., & Stretch, V. (2000). The case against a criterion-shift account of false memory. *Psychological Review, 107*(2), 368-376. doi: 10.1037/0033-295X.107.2.368

Wood, M. J., & Blair, M. R. (2010). Informed inferences of unknown feature values in categorization. *Memory & Cognition*, *39*(4), 666-674. doi: 10.3758/s13421-010-0044-1.

Xu, Y. D. (2002). Feature integration across parts in visual search. Perception, 31(11), 1335-1347. doi: 10.1068/p3431

Yonelinas, A. P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory & Cognition*, *25*(6), 747-763.

APPENDICES

APPENDIX A - INFORMED CONSENT FORM

Gönüllü Katılım Formu

Bu çalışma, Bilişsel Bilimler Bölümü doktora öğrencisi Selda Eren tarafından yürütülen, görsel algı uzerine bir çalışmadır. Çalışmanın amacı, nesne algısı ve nesne hafızası ile ilgili zihinsel süreçler hakkında bilgi toplamaktır. Çalışmaya katılım tamamen gönüllülük temelinde olmalıdır. Sizden kimlik belirleyici hiçbir bilgi istenmemektedir. Cevaplarınız tamamen gizli tutulacak ve sadece araştırmacılar tarafından değerlendirilecektir. Elde edilecek bilgiler bilimsel yayınlarda kullanılacaktır.

Bu deney, fiziksel ve/ya ruhsal sağlığı tehdit edici ya da stres kaynağı olabilecek hiçbir unsur içermemektedir. Ancak, katılım sırasında herhangi bir nedenden ötürü kendinizi rahatsız hissederseniz deneyi yarıda bırakıp çıkmakta serbestsiniz. Böyle bir durumda anketi uygulayan kişiye, deneyi tamamlamadığınızı söylemek yeterli olacaktır. Deney sonunda, bu çalışmayla ilgili sorularınız cevaplanacaktır. Bu çalışmaya katıldığınız için şimdiden teşekkür ederiz. Çalışma hakkında daha fazla bilgi almak için Bilişsel Bilimler Bölümü öğretim üyelerinden Doç. Dr. Annette Hohenberger (Oda: A-219; Tel: 210 3789; E-posta: hohenberger@ii.metu.edu.tr) ya da doktora öğrencisi Selda Eren (E-posta: e115275@metu.edu.tr) ile iletişim kurabilirsiniz.

Bu çalışmaya tamamen gönüllü olarak katılıyorum ve çalışmanın amacı konusunda bilgilendirildim. İstediğim zaman deneyi yarıda kesip çıkabileceğimi biliyorum. Verdiğim bilgilerin bilimsel amaçlı yayınlarda kullanılmasını kabul ediyorum. (Formu doldurup imzaladıktan sonra uygulayıcıya geri veriniz).

İsim Soyad Tarih

İmza

----/----/-----

APPENDIX B - DEBRIEFING FORM

KATILIM SONRASI BİLGİ FORMU

Bu çalışma, Bilişsel Bilimler Bölümü doktora öğrencisi Selda Eren tarafından yürütülen, görsel algı uzerine bir çalışmadır. Nesnelerin algılanması ve hatırlanması sırasında kullanılan nesne gösterimlerinin oluşumu ve aktivasyonu incelenmektedir.

Görsel hafıza literatürü nesne gösterimlerinin renk şekil ve desen gibi özelliklerden oluştuğunu belirtmektedir. Ancak hangi özelliklerin ne şekilde seçildiğine dair yeterli araştırma bulunmamaktadır. Bu çalışmada tercih ettiğimiz yaklaşım ekolojik psikoloji ve algı-eylem alanlarındaki bulguları esas almaktadır.

Test ettiğimiz ilk faktör tekrar eden nesne özelliklerinin kodlanıp kodlanmadığıdır. İkinci sırada ise sekil, renk, desen ve konum gibi farklı nesne özelliklerinin kodlamadaki etkisi incelenecektir.

Bu çalışmadan alınacak verilerin 2010 Kasim ayinda elde edilmesi amaçlanmaktadır. Elde edilen bilgiler sadece bilimsel araştırma ve yazılarda kullanılacaktır. Çalışmanın sonuçlarını öğrenmek ya da bu araştırma hakkında daha fazla bilgi almak için aşağıdaki isimlere başvurabilirsiniz. Bu araştırmaya katıldığınız için çok teşekkür ederiz.

Doç. Dr. Annette Hohenberger (Oda: A-219; Tel: 210 3789; E-posta: hohenberger@ii.metu.edu.tr) Selda Eren (E-posta: e115275@metu.edu.tr)

APPENDIX C – Instructions for Experiment 1

"The experiment consists of two parts. In the first part, you will see a series of slides. There will be objects on these slides. In the second part, I will show you another series of slides and ask you whether you had seen[4] the object during the previous part."

The experimenter opened the Power Point file. "Press spacebar to continue" displayed on black background.

"You will press the spacebar when you are ready to start the first part. You will just watch the slides."

After all 15 slides were displayed, the Power Point turned back to the design view. At that point, the experimenter started the training slides from the beginning and instructed the participants as follows:

"Now I will repeat the same slides for better recall."

After the second round, the experimenter opened the test file, and gave the following instructions: "I will show you a series of slides and ask if you had seen the object in the first part. Reply with Yes or No. Since there is a time limit, try to be as quick as possible."

---

[4] Instead of the standard old/new responses, the participants were required to tell whether they had "seen" the object or not. This is a more appropriate way of asking whether the object is "old" or "new" in Turkish ("gördüm" is the word for "seen" and "görmedim" is the word for "not seen" in Turkish).

APPENDIX D - INFORMED CONSENT FORM

1956

Orta Doğu Teknik Üniversitesi
Middle East Technical University

Öğrenci İşleri Dairesi
Başkanlığı
Registrar's Office

06531 Ankara, Türkiye
Phone: +90 (312) 2103417
Fax: +90 (312) 2107960
www.oidb.metu.edu.tr

B.30.2.ODT.72.00.00 /400-9586-612

08/12/2009

ENFORMATİK ENSTİTÜSÜ MÜDÜRLÜĞÜNE

İlgi: 25/11/2009 tarih ve 706- 14504 sayılı yazınız.

Üniversitemiz Bilişsel Bilimler Anabilim Dalı Doktora öğrencisi Selda Eren'in, 23/11/2009-01/09/2010 tarihleri arasında tezi ile ilgili *"Formation and Activation of Affordance- Based Object Representations: Evidence from Behavirol and Neuroimaging Studies"* başlıklı araştırma çalışmasına ilişkin olarak Orta Doğu Teknik Üniversitesi Psikoloji Bölümü ve ODTÜ Enformatik Enstitüsünde uygulama yapmak için, öğrencinin isteği doğrultusunda görevlendirilmesi Etik Komite onayı ile uygun görülmüştür.

Gereğini bilgilerinize arz ederim.

Saygılarımla.

Nesrin Ünsal
Öğrenci İşleri Dairesi Başkanı

Ekler:
İAEK Başvuru Kontrol Listesi
İAEK Başvuru Formu
İAEK Başvuru Formu Proje Bilgi Formu

BD

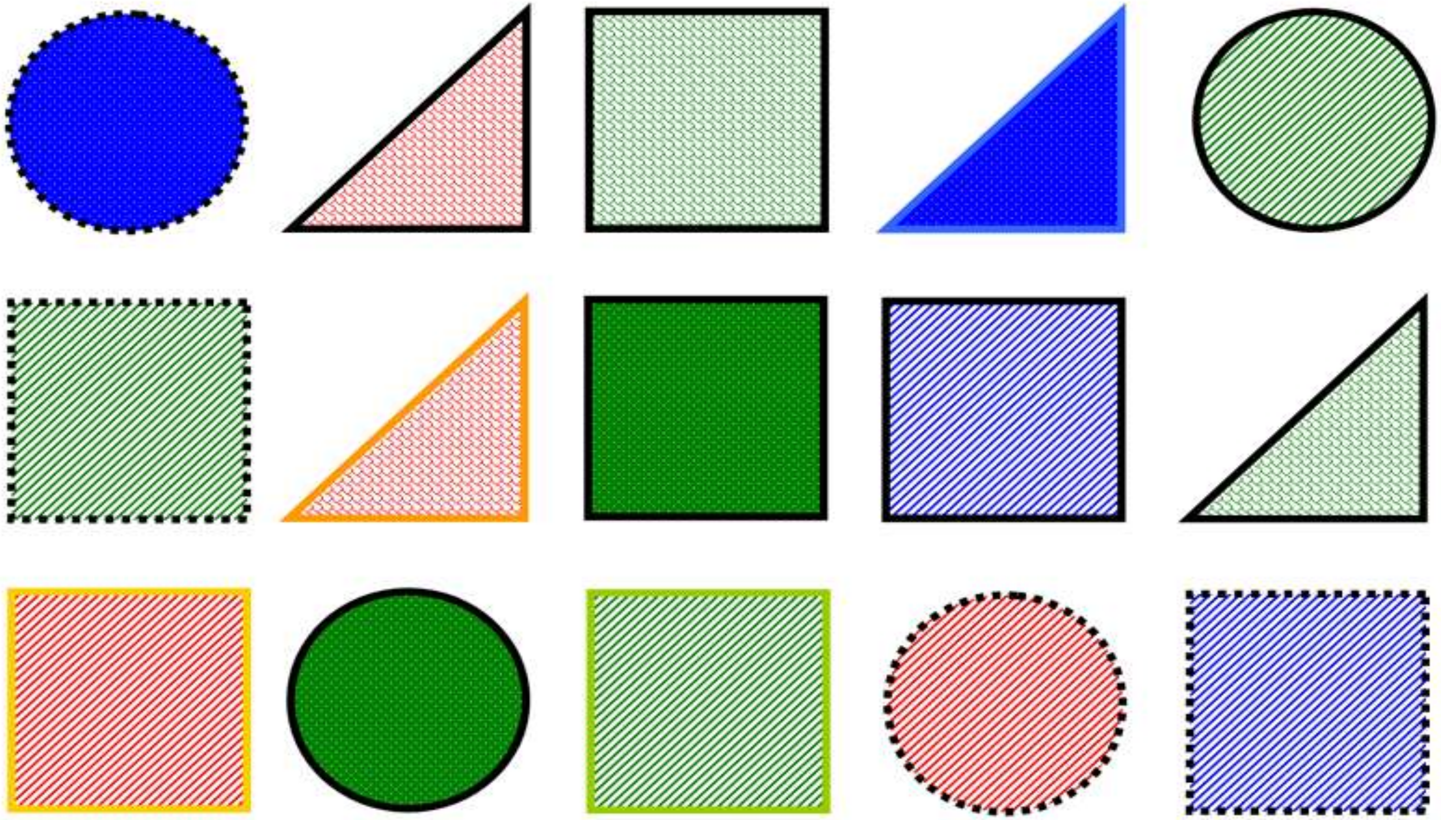
APPENDIX E - CODE FOR THE CDZ-VIS MODEL

```
#include <cstdlib>
#include <iostream>
#include <fstream>
using namespace std;

class CDZ{
public:
   int features[4];
   int layer;
   int weight;
};

int fillarray (CDZ CDZ_array[]) {
   ifstream study ("study.txt");
   int feature;
   int count=0;
   int numfeat=0;

   while ( study.good() )
   {
      for (int i=0;i<4;i++)
      {
         study >> feature;
         CDZ_array[count].features[i]=feature;
      }
      CDZ_array[count].layer=4;
      count++;
   }

   cout << "printing array\n" ;

   for (int j=0;j<count;j++)
   {
      numfeat=CDZ_array[j].layer;
      cout << "CDZ: ";
      for (int i=0;i<numfeat;i++)
         cout << CDZ_array[j].features[i] << " ";
      cout << "layer:" << CDZ_array[j].layer << '\n';
   }

   cout << "end of array\n" ;

   study.close();
   return count;
}

int cleanCDZ(CDZ CDZ_array[],int count,int cdzcount){
   int cdzsim=0;
   int lastcount;
   lastcount=cdzcount; /*keep a copy of cdzcount since it might be changed soon*/
```

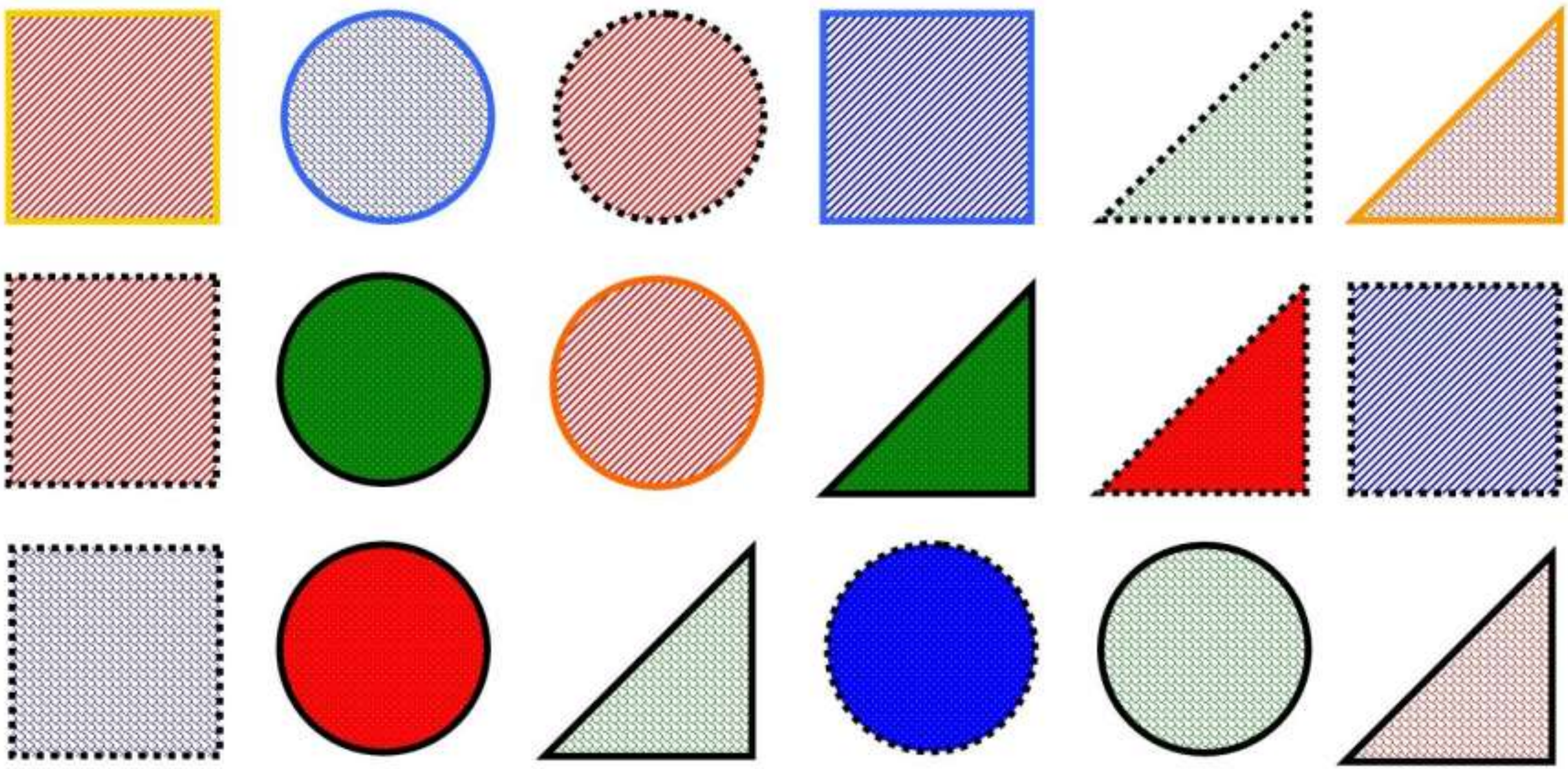
```
   for (int i=count;i<cdzcount-1 && cdzcount==lastcount ;i++) /* until cdzcount-1 not to
compare with itself*/
   {
      if (CDZ_array[i].layer==CDZ_array[cdzcount-1].layer)
      {
         for (int j=0;j<CDZ_array[i].layer;j++)
         {
            if (CDZ_array[i].features[j]==CDZ_array[cdzcount-1].features[j])
               cdzsim++;
         }
      }

      if (cdzsim==CDZ_array[i].layer)
      {
         cdzcount--;
         CDZ_array[i].weight=(CDZ_array[i].weight)+1;
         cout << " burda weight artiyor:" << CDZ_array[i].weight << '\n';
      }
      cdzsim=0;       /* reset cdzsim */
   }
   return cdzcount;
}

int createCDZ (CDZ CDZ_array[],int count,int cdzcount,int objectid ){
   int sim=0;

   for (int j=0;j<objectid;j++)
   {
      /*cout << "alooo\n";*/
      for (int i=0;i<4;i++)
      {
         if (CDZ_array[objectid].features[i]==CDZ_array[j].features[i])
         {
            CDZ_array[cdzcount].features[sim]=CDZ_array[objectid].features[i];
            sim++;
            cout << "benzer f geliyooor" << CDZ_array[objectid].features[i] << "\n";
         }
      }
      if (sim>0)
      {
         CDZ_array[cdzcount].layer=sim;
         CDZ_array[cdzcount].weight=1;
         cdzcount++;
         cdzcount=cleanCDZ(CDZ_array,count,cdzcount);
      }
      sim=0;  /*reset number of similar features  */
   }
   return cdzcount;
}

int CreateNetwork(CDZ CDZ_array[],int count){
   int cdzcount=0;

   cdzcount=count;    /*upper layer CDZs start after layer4 CDZs*/
   for (int i=1;i<count;i++)
   {
```

```
        cdzcount=createCDZ(CDZ_array,count,cdzcount,i);
        cout << "printing array " << i << ":\n";
        for (int j=count;j<cdzcount;j++)
        {
            cout << "yazdirmadan onceki layer info:" << CDZ_array[j].layer << " weight: " <<
CDZ_array[j].weight << "\n";
            for(int k=0;k<CDZ_array[j].layer;k++)
                cout << CDZ_array[j].features[k] << " ";
            cout << "\n";
        }
    }
return cdzcount;
}

int main(int argc, char** argv) {
    int count=0;        /*number of layer 4 CDZs in the array*/
    int cdzcount=0;     /*number of upper layer CDZs in the array*/
    int numfeat=0;
    CDZ neuron[100];

    cout << "Hello World! \n";
    count=fillarray(neuron);
    cdzcount=CreateNetwork(neuron,count);

    for (int j=0;j<cdzcount;j++)
    {
        numfeat=neuron[j].layer;
        cout << "CDZ: ";
        for (int i=0;i<numfeat;i++)
            cout << neuron[j].features[i] << " ";
        cout << "layer:" << neuron[j].layer << " weight:" << neuron[j].weight << '\n';
    }

    return 0;
}
```

APPENDIX F - Objects in the study phase of feature frequency experiment



APPENDIX G - Objects in the test phase of feature frequency experiment



APPENDIX H – ANOVA for the feature frequency experiment

Pair 1 as FRFs: Green color and solid black border

Table 13. Mauchly's test of sphericity

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.
	Greenhouse-Geisser	Huynh-Feldt	Lower-bound	Greenhouse-Geisser
fam	1,000	,000	0	.
feat	,781	4,458	2	,108
fam * feat	,827	3,418	2	,181

Table 14. Tests of within-subjects effects

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
fam	**Sphericity Assumed**	4,800	1	4,800	46,769	,000	,711
	Greenhouse-Geisser	4,800	1,000	4,800	46,769	,000	,711
	Huynh-Feldt	4,800	1,000	4,800	46,769	,000	,711
	Lower-bound	4,800	1,000	4,800	46,769	,000	,711
Error(fam)	**Sphericity Assumed**	1,950	19	,103			
	Greenhouse-Geisser	1,950	19,000	,103			
	Huynh-Feldt	1,950	19,000	,103			
	Lower-bound	1,950	19,000	,103			
feat	**Sphericity Assumed**	1,667	2	,833	13,571	,000	,417
Error(feat)	**Sphericity Assumed**	2,333	38	,061			
fam * feat	**Sphericity Assumed**	,950	2	,475	3,574	,038	,158
Error(fam*feat)	**Sphericity Assumed**	5,050	38	,133			

Pair 2 as FRFs: Square shape and oblique pattern

Table 15. Mauchly's test of sphericity

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.
	Greenhouse-Geisser	Huynh-Feldt	Lower-bound	Greenhouse-Geisser
fam	1,000	,000	0	.
feat	,970	,542	2	,763
fam * feat	,947	,975	2	,614

Table 16. Tests of within-subjects effects.

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
fam	**Sphericity Assumed**	5,002	1	5,002	28,891	,000	,603
	Greenhouse-Geisser	5,002	1,000	5,002	28,891	,000	,603
	Huynh-Feldt	5,002	1,000	5,002	28,891	,000	,603
	Lower-bound	5,002	1,000	5,002	28,891	,000	,603
Error(fam)	**Sphericity Assumed**	3,290	19	,173			
	Greenhouse-Geisser	3,290	19,000	,173			
	Huynh-Feldt	3,290	19,000	,173			
	Lower-bound	3,290	19,000	,173			
feat	**Sphericity Assumed**	,804	2	,402	5,668	,007	,230
Error(feat)	**Sphericity Assumed**	2,696	38	,071			
fam * feat	**Sphericity Assumed**	1,579	2	,790	10,894	,000	,364
Error(fam*feat)	**Sphericity Assumed**	2,754	38	,072			

APPENDIX I – ANOVA for the similarity experiment

Table 17. Tests of within-subjects effects.

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Observed Power(a)
sim	,150	2	,075	2,371	,106	,101	,453
Error(sim)	1,329	42	,032				

Table 18. Tests of within-subjects contrasts

Source	sim	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Observed Power(a)
sim	**Level 2 vs. Level 1**	,003	1	,003	,033	,858	,002	,053
	Level 3 vs. Previous	,223	1	,223	7,364	,013	,260	,735
Error(sim)	**Level 2 vs. Level 1**	1,810	21	,086				
	Level 3 vs. Previous	,636	21	,030				

APPENDIX J – ANOVA for the model comparisons

Feature frequency experiment

Table 19. Tests of within-subjects effects

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Study condition	Sphericity Assumed	1.40	1	1.40	56.76	.000	.37
	Greenhouse-Geisser	1.40	1	1.40	56.76	.000	.37
	Huynh-Feldt	1.40	1	1.40	56.76	.000	.37
	Lower-bound	1.40	1	1.40	56.76	.000	.37
Study condition * model	Sphericity Assumed	1.30	4	.33	13.17	.000	.36
	Greenhouse-Geisser	1.30	4	.33	13.17	.000	.36
	Huynh-Feldt	1.30	4	.33	13.17	.000	.36
	Lower-bound	1.30	4	.33	13.17	.000	.36
Error (Study condition)	Sphericity Assumed	2.35	95	.02			
	Greenhouse-Geisser	2.35	95	.02			
	Huynh-Feldt	2.35	95	.02			
	Lower-bound	2.35	95	.02			

Similarity Experiment

Table 20. Tests of within-subjects effects.

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
oldnew	Sphericity Assumed	.96	1	.96	48.71	.000	.34
	Greenhouse-Geisser	.96	1	.96	48.71	.000	.34
	Huynh-Feldt	.96	1	.96	48.71	.000	.34
	Lower-bound	.96	1	.96	48.71	.000	.34
oldnew * model	Sphericity Assumed	.09	4	.02	1.20	.314	.05
	Greenhouse-Geisser	.09	4	.02	1.20	.314	.05
	Huynh-Feldt	.09	4	.02	1.20	.314	.05
	Lower-bound	.09	4	.02	1.20	.314	.05
Error(oldnew)	Sphericity Assumed	1.87	95	.02			
	Greenhouse-Geisser	1.87	95	.02			
	Huynh-Feldt	1.87	95	.02			
	Lower-bound	1.87	95	.02			

SELDA EREN KANAT

Informatics Institute
Middle East Technical University
06531 Ankara, Turkey
e-mail: seldaeren@gmail.com

EDUCATION

2006 -	**Middle East Technical University**
	Cognitive Science, Ph.D.
2002 - 2006	**Middle East Technical University**
	Information Systems, M.Sc.
2000 - 2002	**Middle East Technical University**
	Mathematics Minor Program
1998 - 2002	**Middle East Technical University**
	Computer Education and Instructional Technology

RESEARCH EXPERIENCE

Year	**Place**	**Enrollment**
2006 - 2008	METU, BAB	Researcher

2003 - 2006	METU, KOVAN	Researcher

WORK EXPERIENCE

Year	**Place**	**Enrollment**
2002 - 2008	METU	Research assistant

PUBLICATIONS

1. Eren-Kanat, S., & Hohenberger, A. (2011). Formation of visual object representations in the convergence-divergence zone framework: An empirical study on the effects of discrete feature similarity and repetition frequency on object familiarity. *Paper presented at the 12th Annual Meeting of the European Congress of Psychology*.
2. Eren, S., & Hohenberger, A. (2010). Feature repetition effects on object familiarity: Evidence from an old/new recognition task. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Meeting of the Cognitive Science Society* (pp. 754-759). Austin, TX: Cognitive Science Society.
3. Eren, S., & Hohenberger, A. (2009). Encoding of affordance-relevant object features affects responses in future contexts. *Proceedings of the 16th Meeting of the European Society for Cognitive Psychology* (p. 12). European Society for Cognitive Psychology.

4. Eren, S. (2008, July). *Object representations in long-term memory as feature sets derived from affordances*. Paper presented at the XXIXth International Congress of Psychology, Berlin, Germany.
5. Eren, S. (2007, August). *Object individuation in Turkish preschoolers and adults: Effects of proximity and alignment*. Paper presented at the 13th European Conference on Developmental Psychology (ECDP), Jena, Germany.
6. Eren, S. (2006). Towards learning affordances: Detection of relevant features and characteristics. *Unpublished Master's Thesis,* Middle East Technical University.

References

Annette Hohenberger
hohenberger@ii.metu.edu.tr
Bilişsel Bilimler Bölümü
Orta Doğu Teknik Üniversitesi
Ankara, 06531 Turkey