

Using Neural Networks in Advertising Research¹

N.J. Stoyanoff
The Galileo Company

Joseph Woelfel
University at Buffalo

Beth Pincura
Johnson Controls
Automotive Systems Group

Paper Presented at the Annual Conference
of the Advertising Research Foundation

April 20, 1994

¹ Paper presented before the Advertising Research Foundation Conference, New York, April, 1994. Trademarks used in this paper are the property of their respective owners. None of the research examples reported represent actual data, but have been constructed and/or modified to protect the proprietary rights of their owners. Results are presented for illustrative purposes only.

Introduction

At many advertising agencies, research plays a fundamental role in the development and evolution of an advertising campaign. When used properly, research can provide insight into marketplace dynamics, and shed light on how consumers are responding to creative materials. The purpose of this document is to illustrate how a relatively new type of computer software, artificial neural networks, can further enhance the role of research by automating rote tasks, and expanding analysis options.

This paper begins with a brief discussion on neural network software, and then reviews a number of specific applications where neural network software has already proven to be a useful technology. Where appropriate, we also discuss some of the theoretical and methodological considerations at the heart of different applications.

Artificial Neural Network Technology

Artificial Neural Networks (ANNs) represent a new class of analytical software. Developed jointly by neuroscientists and software engineers, the algorithms which underlie neural networks attempt to model some of the neural processes which take place in the human brain. While still an emerging technology, the most successful ANNs developed thus far are those that have been optimized to accomplish pattern recognition tasks.

For example, over the last decade the American military has developed several different types of ANNs to help fighter pilots with terrain classification, and weapons targeting. Similarly, many industrial applications have been developed

for managing manufacturing processes and checking the quality of parts. In the financial arena, ANNs have been developed to evaluate risk, and accomplish program trading. By now, thousands of applications of neural technology are in everyday use. Among the areas where neural networks have just recently begun to have great impact is the advertising community. Just as pilots count on neural technology to help them map enemy terrain and guide them to their target, advertisers are discovering how to use neural networks to map the marketplace and guide products and services to their target position.

What is a Neural Network?

However complicated they may become in real life, even the most complex neural network is simply a set of neurons in which each neuron is “connected” to at least some of the other neurons. A *neuron* is basically a device which can be active or inactive.² What we mean when we say that two or more neurons are “connected” is that neurons may “communicate” their activation value across connections to other neurons and thereby influence their activation levels.³

² In more complex neural networks, neurons may not only be active or inactive, but may vary in how active they are. Networks with only neurons which can be active or inactive are often called “binary” networks, while those with neurons which can take on a range of activation values are called “analog” networks. Of the two, analog networks are the more powerful.

³ As is the case with neurons, which can be either binary (active or inactive) or analog (capable of a range of activation), so too connections among neurons may be binary (connected or not connected) or analog (capable of a continuous range of connectivity). As with binary and analog neurons, networks whose connections can take on analog values are immensely more powerful than those which can only take on two discrete values.

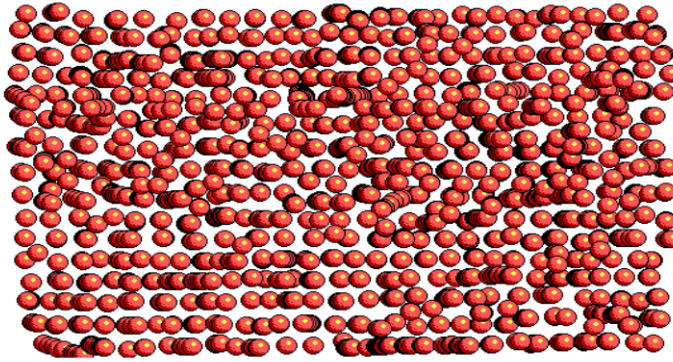


Figure 1: A Slab of Unconnected Neurons

Types of Neural Networks

What a neural network can do is determined entirely by the kinds of neurons it has, and the way in which they are connected, which, together are called the *architecture* of the network. Two major types of network are commonly mentioned, *supervised* (usually backpropagation) models, and *unsupervised* or *self organizing* models. Of the two, the supervised models are the older, and perhaps the least interesting model. Supervised neural networks work much like regression models, although they may offer some important advantages over the classic regression equation under certain circumstances. Supervised models have also been more widely described than unsupervised models, and an extensive literature about supervised models is already available. (Woelfel, 1993; McClelland, et. al., 1988) In this paper we will focus more sharply on unsupervised networks, which are less well known in the advertising community, and which seem to offer more promise for future applications than the better known supervised models.

Self Organizing Neural Networks

A self organizing neural network is a network in with a series of neurons (either binary or analog) , which may or may not be connected to each other in any particular way. When exposed to stimuli from the environment, some of the neurons become active. So, for example, a neural network might be exposed to a cat, which will result in turning on certain neurons in the form of a cat.⁴ Patterns of neuron activation represent a picture or pattern in a way analogous to a theater marquee.

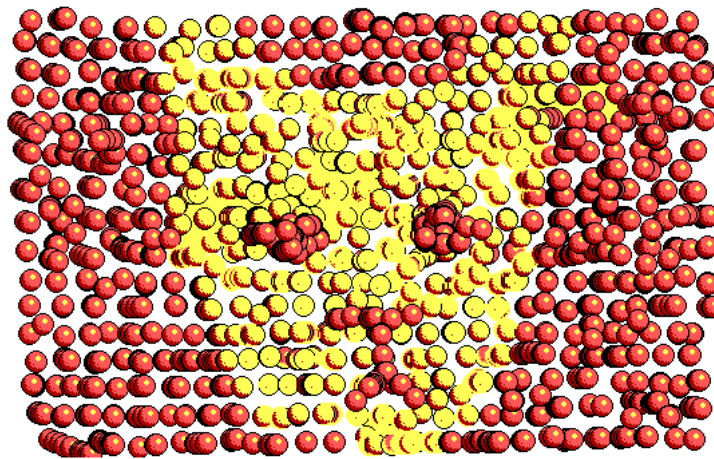
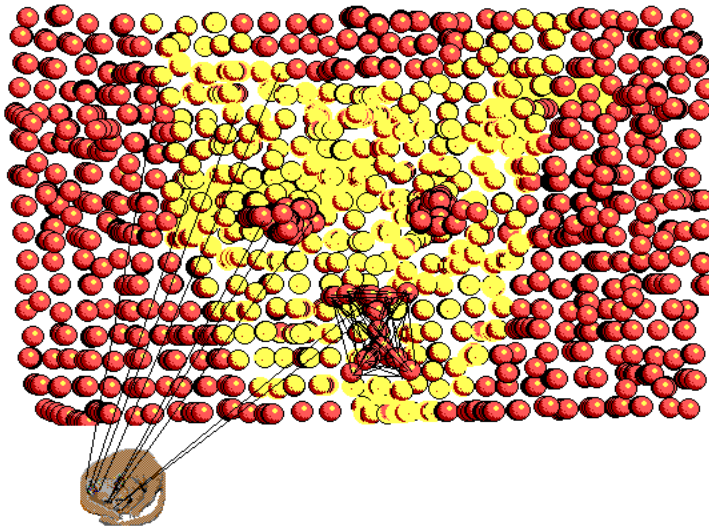


Figure 2: Slab of Neurons Representing a Cat

Learning

⁴ In a human brain, perhaps a dozen separate images are formed in the visual cortex at the back of the brain.

A self-organizing neural network can not only represent (display) a pattern, but it can remember it as well. Whenever two or more neurons are simultaneously active, they tend to become increasingly tightly connected to each other. These connections (*synapses* in natural networks) are the network's memory. Figure 3 shows how glancing quickly at a cat may give one a partial picture; in this example, the viewer perceives an ear, an eye, and a nose of a cat.



Since the image of the cat consists of *interconnected* neurons, (only the interconnections in the nose have been drawn in the Figure to avoid too much clutter) exciting a sufficient subset can be enough to communicate the activation to the rest of the neurons in the pattern, and one can indeed remember the entire cat.

Multiple Patterns

Just as the self-organizing network can connect parts of a pattern to other parts of the same pattern, it can connect one pattern to others. In this way the picture of the cat could be

connected to the sound “meow” or the word “C”, “A”, “T”. Taken together, these patterns represent the linkages familiar from “free association” studies. It’s easy to see how understanding the way in which neural networks represent, store, link and retrieve patterns can be interesting to advertising researchers.

Developing Marketing and Advertising Applications

Using neural networks to accomplish advertising research has appeal for 3 reasons. First, the fundamental task behind all data analysis performed by an advertising researcher is pattern recognition. That is, a sample is drawn; questions are asked; and the analyst’s job is to see if there is some pattern to the responses to each question. As mentioned above, ANNS have proven to be particularly good at pattern recognition tasks. In short, the basic task is the same, and there is no reason to believe that the phenomena of interest to advertisers couldn’t be similarly studied using this technology.

Second, advertisers either study cognitive processes directly (e.g., opinion formation and attitude change over time) or the relationship between cognitive processes and behavior. Because ANNs utilize algorithms that are modeled after the biological processes that underlie cognitive processes, there is a certain heuristic appeal of using an analytic technology with such a close ties to the phenomena under study.

Third, ANNs have been shown to be particularly useful for intuitive, “gut-level” analysis, such as recognizing a familiar face, or guessing whether a loan applicant will prove reliable or a deadbeat. Neural networks can discover patterns not only in numerical data, but text as well. They can portray the holistic impression respondents get when they view an advertisement or see a story board. They can follow the chain of associations evoked by a product or a particular appeal.

Available Software

While there are a plethora of neural network programs currently available in the commercial arena, very few have been specifically designed for use by the advertising community. The notable exceptions are the suite of programs developed to date by The Galileo Company and, in particular, a program that has been optimized to read and analyze text -- CATPAC.

CATPAC

CATPAC was written, quite simply, *to read an existing text and report the main ideas in that text*. CATPAC begins by running a *scanning window* (default window size is seven words) through the text, as in the following example:

Detroit is exciting! You can feel the excitement of it. It's big, and brawling.
You have to watch your hat and coat, but its worth it.
You can have a good time in
Detroit.

There's lots of things to do in Detroit. Detroit is a fun city. Lots of driving;
Buffalo is a friendly city. They call it the city of good neighbors, and that's true. Buffalo is the friendliest city. Buffalo is a working town, but the people
a party there.

The highlighted words are in the seven word scanning window.⁵ Each word in the text is assigned a neuron in

⁵ Although it seems there are 12 words in the window, CATPAC routinely ignores words like *is, the of, and it*.

CATPAC's artificial "brain", and, when a word is in the scanning window, its corresponding neuron is activated. The window will then scan one word to the right, which means the first word will leave the window and the eighth word will enter. This process continues until all words have been read.

Each time the window scans, the neurons in it are activated, and the connections among neurons that are simultaneously active are strengthened. As a result, words which occur relatively close together in the text will end up strongly connected, while those further apart will be less strongly connected.⁶ This procedure differs from simple cooccurrence models in two ways: first, while the neurons which represent words in a present window will always be active, other neurons may also be active because they are connected to words in the window. Secondly, indirect relations are completely accounted for in the CATPAC model, since two words which never cooccur in the window may still become related because they are related to another word or words which is related to both of them.

The main result of a CATPAC analysis is a complete word count, a square matrix of word similarities (the connection weights), a diameter method cluster analysis which shows the main clusters of ideas in the text, and a perceptual map which arrays the most frequently occurring words in a highly visual display.

Using CATPAC for Copy Test Research

While copy testing methodologies can vary, most ultimately report on 3 key issues:

⁶ Due to normalization rules, the least proxemic words will be negatively connected.

1. recall of the specific product (or service) or advertised,
2. effective communication of the “main idea” contained in the ad, and
3. the relative persuasiveness of the ad to stimulate purchase interest.

Studying Recall Measures.

Perhaps the most fundamental question a copy test can answer is... can someone who views a commercial remember the name of the product or service advertised? Many advertising theorists argue that accurate recall is essential to advertising effectiveness. Almost always, the question is put quite straightforwardly...

What was the name of the product (or service) in the ad you just saw?

The simplest output CATPAC can provide is a simple word count. In fact, regardless of the type of analysis you perform with CATPAC, the output will always contain a listing of the words contained in the text file, and a count of how frequently each word occurs.

Exhibit 1 shows a set of 10 responses one might see after copy testing an ad for a Chevy S10 truck, and asking people to name the product or service advertised. Note: the -1 marker in the data is file simply

Exhibit 1	
Responses to the	
Question:	
What was the name of	
the product or service	
you just saw an ad for?	
GM Trucks	
-1	
General Motors	
-1	
Chevy	
-1	
Chevrolet	
-1	
Chevy Trucks	
-1	
Chevrolet	
-1	
Chevy S10	
-1	
S10 Truck	
-1	
S10	
-1	
Chevy S10	
-1	

a marker which delimits each person's response.

As illustrated in Exhibit 1, when responding to a "recall" type question, most respondents will correctly identify the product. Occasionally, however, some respondents will respond with an abbreviated name, or the name of the parent company which markets the product or service. The percentages of such responses are usually of interest to the analyst.

Exhibit 2 shows a CATPAC analysis of this data set, and details the frequency with which each word was mentioned. The left-most column shows a list of each word read by the program in descending order of frequency. Following the raw frequency counts, there are 2 "percentage" columns.

The first percentage column is obtained by dividing the word frequency by the total of all words. Hence, if there were 16 total unique words in the original file, and a certain word occurs 4 times, the percentage is 25%.

The second percentage column is obtained by noting the number of episodes (out of all episodes) containing a certain response occurs. So, for example, if there are 10 episodes, and a certain word (or unique response) occurs in 4 episodes, the resulting percentage is 40%.

The data in the right-most columns is exactly the same as what was in the left-most columns, but now the words are listed in alphabetical order for easy look-up.

Main Idea Communication. Beyond recall, the next issue

Exhibit 2
CATPAC Analysis of the data in Exhibit 1

CATPAC_PC v3.00
03/28/94 23:13:31

TITLE: Catpac Analysis of Chevy Truck Data
DATA FILE: Data shown in Exhibit 1

TOTAL WORDS 16
TOTAL UNIQUE WORDS 8
TOTAL EPISODES 10
TOTAL LINES 10

DESCENDING FREQUENCY LIST	ALPHABETICALLY SORTED LIST
---------------------------	----------------------------

WORD												
	FREQ	PCNT	CASE	CASE		FREQ	PCNT	CASE	CASE		FREQ	PCNT
CHEVY	4	25.0	4	40.0	CHEVROLET	2	12.5	2	20.0			
S10	4	25.0	4	40.0	CHEVY	4	25.0	4	40.0			
TRUCKS	2	12.5	2	20.0	GENERAL	1	6.3	1	10.0			
CHEVROLET	2	12.5	2	20.0	GM	1	6.3	1	10.0			
GM	1	6.3	1	10.0	MOTORS	1	6.3	1	10.0			
GENERAL	1	6.3	1	10.0	S10	4	25.0	4	40.0			
MOTORS	1	6.3	1	10.0	TRUCK	1	6.3	1	10.0			
TRUCK	1	6.3	1	10.0	TRUCKS	2	12.5	2	20.0			

typically examined during copy testing is the main idea registration. Here, the goal is to try and ascertain whether or not people viewing an ad understood its main message. As

with the recall question, typically, the main idea question is usually posed in a very straight forward manner...

Exhibit 3
Responses to the Question:
What was the main point of the ad
you just saw for the new Ford
Mustang?

Its the car of the year
-1
Buy a Mustang
-1
It is what it was
-1
Buy a Mustang
-1
Buy a Mustang, its the car of the year
-1
Buy a Mustang, it is what it was
-1
Buy Mustang
-1
Mustang is the car of the year
-1
Buy a Ford Mustang
-1
Buy a Mustang
-1

What was the main point of
the commercial you just
saw?

Exhibit 3 shows a set of responses one might see in response to a recent ad for the Ford Mustang (see Appendix A to review the storyboard for the

commercial to which we are referring).

Exhibit 4 shows the results of a CATPAC analysis of the data presented in Exhibit 3. Exhibit 4 first shows a listing of the words contained in the data file. Then, below the word counts, there is a figure called a dendogram. A dendogram is the one of the outputs produced by an analytic procedure called a *hierarchical cluster analysis*.

A dendogram contains two important features: (1) a list of the words contained in the data file, and (2) a series of markers that denote word associations. In Exhibit, the marker is a caret mark like this... ^

In Exhibit 4, the way in which the caret marks are typed beneath the words slightly resembles the skyline of a city as seen from afar. Within this pseudo-skyline, each building represents a series of words which frequently co-occurred within the original text file, and in so doing, represents the main ideas contained in the text file.

A close inspection of the dendogram in Exhibit 4 shows 3 main groupings (or clusters). The first cluster contains the words (ITS) CAR OF THE YEAR; the second cluster contains the words IS WHAT IT WAS; and the third cluster contains the words MUSTANG BUY A FORD.

When referencing this output (Exhibit 4) against the original data file (Exhibit 3), we can see that CATPAC did a fairly good job of capturing the responses to the question.

Exhibit 4
CATPAC Analysis of the data in Exhibit 3

CATPAC_PC v2.00
 03/31/94 07:09:23

TITLE: Main idea of the new Mustang ad
 DATA FILE: Data shown in Exhibit 5

TOTAL WORDS 50
 TOTAL UNIQUE WORDS 13
 TOTAL EPISODES 10
 TOTAL LINES 10

DESCENDING FREQUENCY LIST

ALPHABETICALLY SORTED LIST

WORD	FREQ		PCNT		WORD	FREQ		PCNT	
	FREQ	PCNT	FREQ	PCNT		FREQ	PCNT	FREQ	PCNT
MUSTANG	8	16.0	8	80.0	A	6	12.0	6	60.0
BUY	7	14.0	7	70.0	BUY	7	14.0	7	70.0
THE	6	12.0	3	30.0	CAR	3	6.0	3	30.0
A	6	12.0	6	60.0	FORD	1	2.0	1	10.0
IT	4	8.0	2	20.0	IS	3	6.0	3	30.0
CAR	3	6.0	3	30.0	IT	4	8.0	2	20.0
OF	3	6.0	3	30.0	ITS	2	4.0	2	20.0
YEAR	3	6.0	3	30.0	MUSTANG	8	16.0	8	80.0
IS	3	6.0	3	30.0	OF	3	6.0	3	30.0
ITS	2	4.0	2	20.0	THE	6	12.0	3	30.0
WHAT	2	4.0	2	20.0	WAS	2	4.0	2	20.0
WAS	2	4.0	2	20.0	WHAT	2	4.0	2	20.0
FORD	1	2.0	1	10.0	YEAR	3	6.0	3	30.0

16 – Neural Networks in Advertising



Exhibit 5
CATPAC Analysis of respondents hearing an HCI presentation

DESCENDING FREQUENCY LIST					ALPHABETICALLY SORTED LIST				
WORD	FREQ	PCNT	CASE FREQ	CASE PCNT	WORD	FREQ	PCNT	CASE FREQ	CASE PCNT
GUNS	17	23.0	56	82.4	17	1	1.4	7	10.3
DANGEROUS	6	8.1	34	50.0	ASSAULT	1	1.4	7	10.3
HAND	6	8.1	29	42.6	BAN	1	1.4	7	10.3
BANNED	4	5.4	22	32.4	BANNED	4	5.4	22	32.4
I	3	4.1	21	30.9	BELIEVE	2	2.7	14	20.6
RID	3	4.1	16	23.5	CHILDREN	1	1.4	7	10.3
CRIMINALS	3	4.1	15	22.1	CRIME	3	4.1	11	16.2
CRIME	3	4.1	11	16.2	CRIMINALS	3	4.1	15	22.1
RIFLES	2	2.7	14	20.6	DANGEROUS	6	8.1	34	50.0
NEED	2	2.7	14	20.6	ELIMINATE	1	1.4	3	4.4
BELIEVE	2	2.7	14	20.6	GUNS	17	23.0	56	82.4
KILL	2	2.7	14	20.6	HAND	6	8.1	29	42.6
PEOPLE	2	2.7	14	20.6	HANDGUNS	2	2.7	14	20.6
HANDGUNS	2	2.7	14	20.6	HUNTER	1	1.4	7	10.3
TOGETHER	2	2.7	12	17.6	I	3	4.1	21	30.9
PARTICULARLY	1	1.4	3	4.4	IMMEDIATELY	1	1.4	7	10.3
THINK	1	1.4	7	10.3	KILL	2	2.7	14	20.6
WE	1	1.4	7	10.3	NEED	2	2.7	14	20.6
BAN	1	1.4	7	10.3	PARTICULARLY	1	1.4	3	4.4
ASSAULT	1	1.4	7	10.3	PEOPLE	2	2.7	14	20.6
IMMEDIATELY	1	1.4	7	10.3	REDUCE	1	1.4	5	7.4
HUNTER	1	1.4	7	10.3	RID	3	4.1	16	23.5
17	1	1.4	7	10.3	RIFLES	2	2.7	14	20.6
SHOTS	1	1.4	7	10.3	S	1	1.4	7	10.3
SPORTSMAN	1	1.4	7	10.3	SHOTGUNS	1	1.4	7	10.3
SHOTGUNS	1	1.4	7	10.3	SHOTS	1	1.4	7	10.3
S	1	1.4	7	10.3	SPORTSMAN	1	1.4	7	10.3
CHILDREN	1	1.4	7	10.3	THINK	1	1.4	7	10.3
REDUCE	1	1.4	5	7.4	TOGETHER	2	2.7	12	17.6
ELIMINATE	1	1.4	3	4.4	WE	1	1.4	7	10.3

18 – Neural Networks in Advertising

Exhibit 6
CATPAC of responses of people who heard NRA message

DESCENDING FREQUENCY LIST					ALPHABETICALLY SORTED LIST				
WORD	FREQ	PCNT	CASE FREQ	CASE PCNT	WORD	FREQ	PCNT	CASE FREQ	CASE PCNT
GUNS	11	18.6	39	73.6	AMENDMENT	1	1.7	7	13.2
GUN	8	13.6	37	69.8	AMILY	1	1.7	7	13.2
LIFE	7	11.9	41	77.4	ATTACKED	1	1.7	7	13.2
SAVE	5	8.5	28	52.8	CHILDREN	2	3.4	11	20.8
PROTECT	4	6.8	19	35.8	CITIZENS	1	1.7	3	5.7
GOOD	3	5.1	11	20.8	DEPEND	1	1.7	7	13.2
RIGHT	3	5.1	15	28.3	FAMILY	1	1.7	7	13.2
HAND	2	3.4	11	20.8	FAVOR	1	1.7	5	9.4
CHILDREN	2	3.4	11	20.8	GOOD	3	5.1	11	20.8
CITIZENS	1	1.7	3	5.7	GUN	8	13.6	37	69.8
FAVOR	1	1.7	5	9.4	GUNS	11	18.6	39	73.6
SECOND	1	1.7	6	11.3	HAND	2	3.4	11	20.8
AMENDMENT	1	1.7	7	13.2	LIFE	7	11.9	41	77.4
RIGHTS	1	1.7	7	13.2	NEED	1	1.7	7	13.2
ATTACKED	1	1.7	7	13.2	NEEDS	1	1.7	7	13.2
YOURSELF	1	1.7	7	13.2	PEOPLE	1	1.7	7	13.2
PEOPLE	1	1.7	7	13.2	PROTECT	4	6.8	19	35.8
DEPEND	1	1.7	7	13.2	PROTECTED	1	1.7	7	13.2
AMILY	1	1.7	7	13.2	RIGHT	3	5.1	15	28.3
NEEDS	1	1.7	7	13.2	RIGHTS	1	1.7	7	13.2
FAMILY	1	1.7	7	13.2	SAVE	5	8.5	28	52.8
PROTECTED	1	1.7	7	13.2	SECOND	1	1.7	6	11.3
NEED	1	1.7	7	13.2	YOURSELF	1	1.7	7	13.2

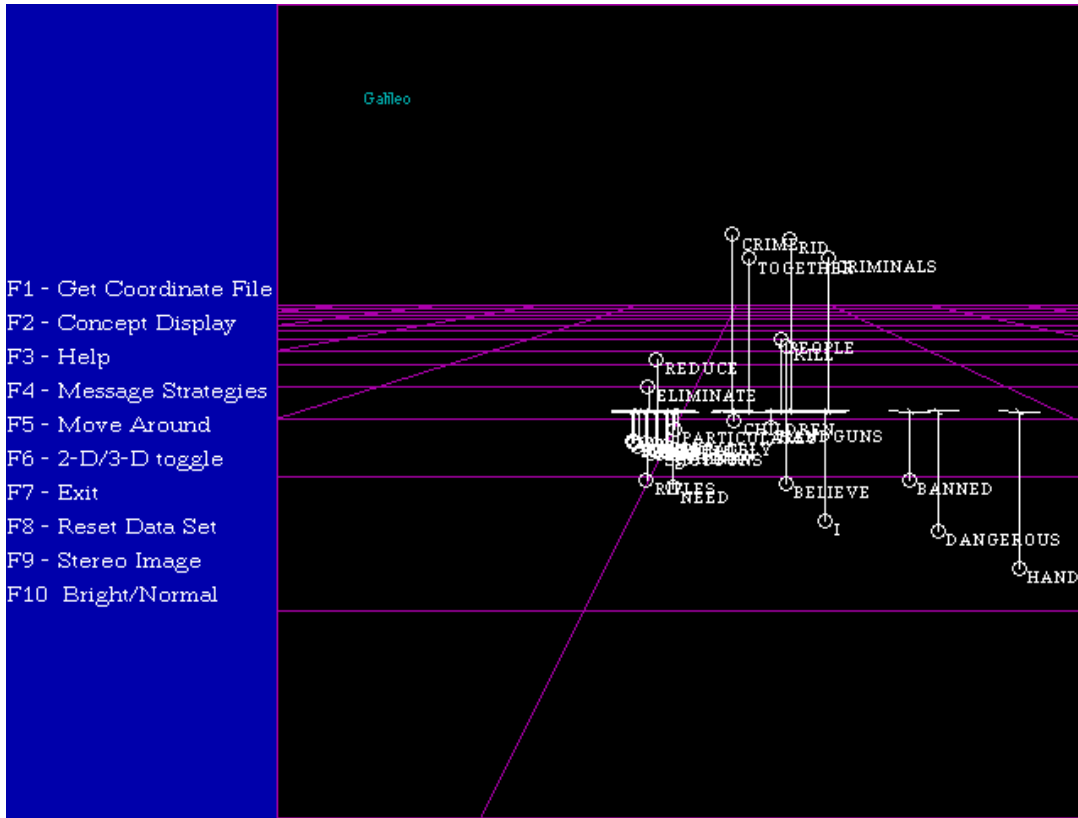


Figure 4: Perceptions of those hearing an anti-gun message

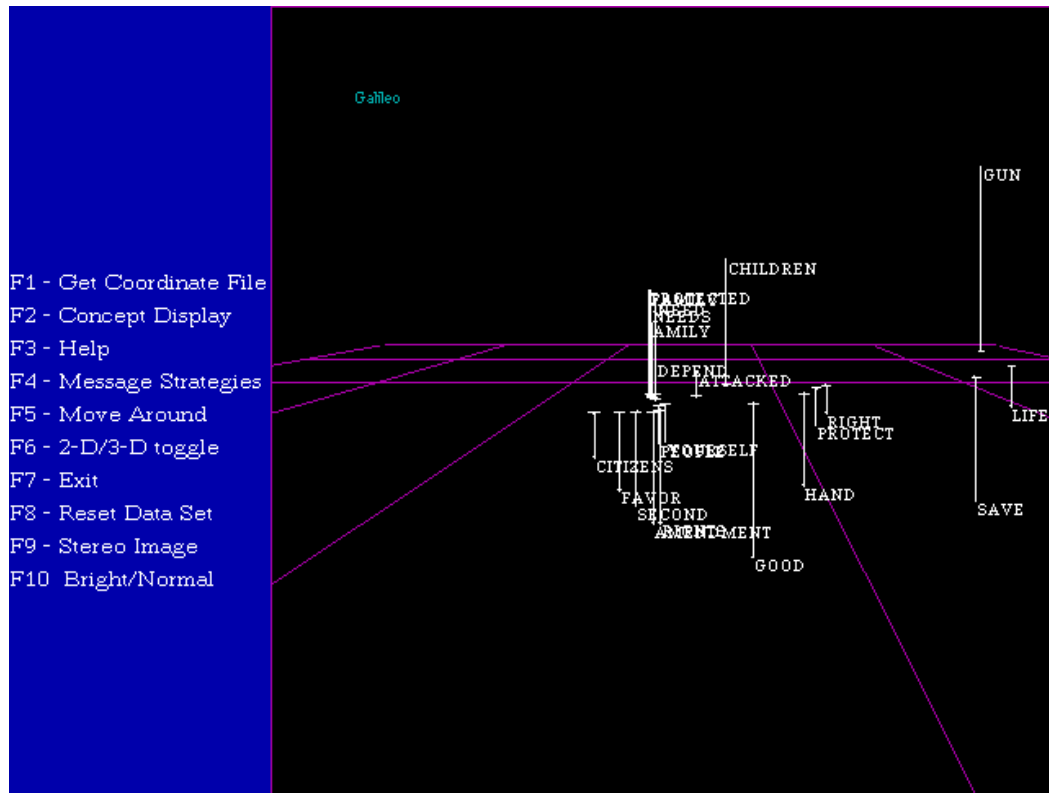


Figure 5: Perceptions of those hearing a pro-gun message

Summary:

o review, CATPAC works by considering each word within a text file as a neuron that initially, has the potential to be connected to every other word (or neuron) contained in the text file. As the program “reads” through the text file, the program adjusts the connection strength between neurons. After reading the entire text file, the result is a word-by-word matrix of connection strengths. This matrix is analogous to a correlation matrix in that its cell entries represent the relative association between words. Inasmuch, this matrix can be decomposed (i.e., factored) using traditional techniques. Then, the first 2 (or 3) rows and columns of the resulting eigenstructure can then be plotted on a standard grid.

