RICE UNIVERSITY

EFFECTS OF REPETITION OF IDENTICAL AND SIMILAR ADVERTISEMENTS ON RETENTION AND ATTITUDE CHANGE

by

Cornelius D. Valenti

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

Thesis Director's Signature:

Houston, Texas

September, 1975
Abstract

EFFECTS OF REPETITION OF IDENTICAL AND SIMILAR ADVERTISEMENTS ON RETENTION AND ATTITUDE CHANGE

by

Cornelius D. Valenti

It is generally held that repetition enhances the effectiveness of television advertisement in terms of information retention and affective rating. The present experiment examined this effect in light of similar predictions made by the encoding variability and mere exposure hypotheses. Seven different groups of 17 subjects each served in a between groups design. Conditions were defined by different levels of presentation frequency (one, three, and six) and repetition similarity (low, medium, and high). The results suggest that greater frequency and lower similarity (higher variability) of repetition enhance retention measures. No firm conclusions could be made concerning the affective ratings. These findings were interpreted as providing support for the encoding variability hypothesis, but not for the mere exposure hypothesis. The implication of these findings for advertising repetition, in particular, and stimulus repetition, in general, were discussed.
Acknowledgments

The author wishes to express sincere appreciation to the members of his thesis committee; Dr. William Howell, who served as Chairperson, and Drs. John Brelsford and Diana Rathjen. Gratitude is extended to Dr. Howell, in particular, who provided considerable guidance to the designing, conducting, and writing of this thesis.

The author would also like to acknowledge the debt owed to Bell Telephone Laboratories, Technical Organization 9131 for funding the academic year and the thesis research at Rice University.

Finally, the author would like to express his thanks to Mrs. Peggy Shaw, who typed the final copy, and his most sincere gratitude to his wife, Mary Beth Valenti, for her enormous clerical assistance and, above all else, moral support without which this thesis would not have been possible.
<table>
<thead>
<tr>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title Page</td>
<td>i</td>
</tr>
<tr>
<td>Abstract</td>
<td>i</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>ii</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>iii</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Method</td>
<td>36</td>
</tr>
<tr>
<td>Results</td>
<td>43</td>
</tr>
<tr>
<td>Discussion</td>
<td>54</td>
</tr>
<tr>
<td>Reference Note</td>
<td>65</td>
</tr>
<tr>
<td>References</td>
<td>66</td>
</tr>
<tr>
<td>Appendix A</td>
<td>71</td>
</tr>
<tr>
<td>Appendix B</td>
<td>75</td>
</tr>
</tbody>
</table>
Introduction

One of the most basic assumptions in psychology for some time has been that organisms are capable of reacting to the repetitive or frequentistic nature of a stimulus. Recently, efforts have been made to provide an explicit account of the processes governing frequency representation and storage. For example, it has been suggested by various theorists that multiple occurrences of a stimulus are stored as increments of a single memory trace (trace-strength hypothesis), multiple and distinguishable traces (multiple-trace hypothesis), or some hybrid of the two (multiple-process hypothesis) (Howell, 1973).

There are also a number of very practical implications of repetition effects. One such practical area is advertising. Ten years ago it was reported that $20 billion was spent annually on advertising in the United States (Alsbaugh, 1964). A large portion of that expenditure was for the repetition of ads. In fact, as expressed by Bass and Barrett (1972, p. 533), there is one tenet which is firmly adhered to by the majority of advertisers--repeat advertising leads to real gains in advertising effect. One need only watch a one or two hour television program sponsored by a single company to observe this principle in practice.
Advertisers usually explain the presumed benefits of repeat advertising in one of two ways: in terms of retention maintenance or of "reach". In the first case, it is held that recall of the advertising will decline rapidly unless it is frequently reinforced by repeated presentations of the ad. In the second case, it is suggested that the message may not even reach the consumer unless it occurs when he is in the market for the product being advertised. Repeat advertising increases the probability that an ad will be present when the consumer is ready to buy. Krugman (1972) relates these two explanations to differing views on the nature of learning. In the former case, the ad must be learned in the same way as a habit—through gradual increments in "strength" with practice. In the latter case, learning may require very few "trials" if the conditions are "ripe" (i.e., it occurs in a "one-trial" fashion). Despite these divergent viewpoints, the implications of both for advertising are similar—repetition should enhance the effectiveness of advertisements as measured by information retention.

Some of the earliest work investigating the effects of repetitive advertising was reported by Strong (1914). He formulated a cube root function for the relationship between retention and identical ad repetition: i.e., the value of the second exposure was hypothesized to be 26% greater than that of the initial presentation, the third, 44% greater than
the first, and so on for subsequent repetitions. From this initial oversimplified and inadequate formulation, further study has revealed that the relationship is far more complex. For example, a study by Stewart (1964) identified a number of variables which moderate Strong's relationship. Stewart found that different products produced different repetition-recall functions. In addition, he observed that some control groups (no ad exposure) exhibited an increase in product awareness as great as that shown by experimental groups. In fact, he even found that one product was purchased more often by control groups than by experimental groups. In sum, repetition of an ad does not guarantee positive retention or marketing functions—cubic or otherwise.

More recent research has yielded results which are also incompatible with the shape of the repetition-recall function described by Strong. Simon (1969) reports that there are no "economies of scale" in advertising. That is, repeated advertisements add less and less to the total retention of the material—a sort of diminishing marginal return. Others raise an even more complicated issue. The results of advertising must be evaluated not only in terms of incremental effects, but also in terms of their effects along a time dimension (Ostheimer, 1970; Ray, Sawyer, and Strong, 1971). In other words, the worth of the nth
repetition is not a set value, but varies according to the time at which it is presented. Finally, there are a host of other potential moderating variables such as the color, size and format of the ad; the medium in which it appears, and the carrying material.

The many potential moderating variables are not the only factors which complicate the effects of ad repetition. Another problem involves the dependent variables in terms of which "effects" are defined. Many studies are concerned only with what viewers remember about ads as measured by recognition or recall scores. An equally important dependent variable is the affect toward the ad or the product. If people retain substantial information from a commercial, but also have strong negative feelings about it, then it would seem possible, at least, that the advertising could do more harm than good. Of course, the ultimate criterion of ad effectiveness is whether the viewer is more likely to buy the product than someone who has not seen the advertising. An approximation to this criterion is found in self-report "intention to buy" measures. Actual buying behavior is extremely difficult to measure as it relates to prior advertising (Britt, 1966).

Another consideration in predicting the effects of ad repetition is the nature of what is considered a repetition. Obviously, some repetitions are more similar than others.
Some, in fact, are identical. There is some evidence, and considerable intuitive appeal, for the notion that degree of similarity influences the viewer's response.

At this point it appears safe to say that there are many variables to be taken into account when attempting to describe the effects of repeat advertising. In general, they can be summarized as: (1) the criterion of effectiveness (retention scores, attitude ratings, intention reports), (2) the definition of a repetition (similarity of different presentations), and (3) the moderating factors (contextual variables, spacing repetitions, and so on).

It is no surprise that advertisers have been interested in the above issues and have studied all of them to one extent or another. The major problem is that the resulting data are not at all clear-cut. In order to gain a better understanding of how ad repetitions might affect consumers, it would perhaps be well to review what is generally known about the cognitive and affective responses of humans to repetitive stimuli. Much of the information bearing on this problem has come from basic research on verbal learning and memory processes, "mere exposure" effects, and persuasive communications. Studies of verbal learning and memory have typically employed simple stimuli, laboratory settings, and information retention responses. Those involving "mere exposure" effects and persuasive communications have em-
ployed both simple and complex stimuli, both laboratory and real world settings, and affective responses.

Repetition and Retention

In the multi-trial free recall paradigm, a list of items is presented to the subject a certain number of times and the subject is then asked to recall as many of the items as possible in any order. One of the most basic empirical findings from studies employing such a task is that the acquisition curve is negatively accelerated or exponential (Murdock, 1974, ch. 8). That is, recall as a function of number of presentations of the entire list resembles a standard learning curve. Waugh (1962) reports single-trial free recall research in which items were repeated within a single list. She found that the probability of recall of an item is a linear function of its presentation frequency.

Hellyer (1962) studied the recall of consonant tri-grams presented with a frequency of one, two, four, or eight presentations (length of retention interval was also manipulated). He found that the recall performance bore a strong positive relationship to frequency for all levels of the retention interval variable. The shape of the recall-frequency relationship was generally linear.

Underwood (1969) conducted a series of studies, primarily on the effects of lag or spacing on free recall,
which also provide information on the repetition effect. He repeated two-syllable nouns from one to four times in a single-trial, free recall paradigm and obtained a linear relationship (similar to Hellyer's) between percent recall and frequency of stimulus presentation. Underwood also found that spacing of repetitions enhanced free recall. Others have reported similar effects of spaced repetitions in a wide variety of experimental situations (Hintzman, 1974).

Instead of citing further examples, it is sufficient to conclude that:

Except for the paradoxical word-frequency effect, in which recognition memory for infrequent words is superior to that for frequent words (e.g. Shepard, 1967), each study merely reinforces the conclusion that learning and retention tend to improve with frequency. (Howell, 1973, p. 46)

The above studies are an indication of the general findings in verbal learning research on the repetition effect for recall. The research can certainly lend insight into the possible repetition effects of commercials. The Underwood and Waugh studies, for example, used the same type of paradigm (single trial, items repeated within a list, recall measures) as many advertising studies. Because retention of advertising is still basically a verbal learning task, though more complex than most, these basic findings suggest that recall of advertising should increase
as a function of frequency—probably in a near linear or negatively accelerated fashion. The retention should also be moderated by such factors as lag (spacing) and response delay.

As noted earlier, the advertising research on repetition is somewhat equivocal. Ray, Sawyer, and Strong (1971) define the "repetition function" for advertising as "the level, shape, and slope of the relationship between repetitive consumer exposure to advertising and the effect of those exposures" (p. 14). The typical repetition function for retention discussed in various media models seldom varies from the negatively accelerated learning curve (Ray, Sawyer, and Strong, 1971). Such a negatively accelerated function fits in very well with the idea of diminishing marginal returns espoused by Simon. It also represents further contradiction to the belief that repeated ads show continually increasing effectiveness in the form of positively accelerated or sigmoid functions (Simon, 1969; 1965). Numerous studies are cited in both of these articles to support Simon's contention.

It is unlikely, however, that the repetition function for retention of ads is always as simple as the linear or negatively accelerated function characteristic of laboratory-based research. It has been observed, for example that as frequency increases beyond the point at which recall is
highest, further stimulus presentations produce a recall decrement. That is, the "extended" repetition function for recall may be of an inverted U-shape (Greenberg and Suttoni, 1973). Since, as we have just seen, learning and retention are usually represented as an increasing function of repetitions, advertising repetition must involve more than these relatively simple processes. Advertisers have suggested three possible complicating factors: attention, decay, and spacing.

Grass (1968) and Grass and Wallace (1969) argue the attention hypothesis as follows:

When a viewer is repeatedly exposed to a TV commercial, attention first increases (generation) to some maximum value (the satiation point or onset of satiation) after which it declines to some equilibrium level which is a function of the frequency or rate of exposure. (Grass and Wallace, 1969, p. 3)

In this formulation, learning is dependent upon attention. After the onset of satiation there is no attention; consequently, there is no new information for the viewer to learn and overall recall declines. Greenberg and Suttoni (1973) provide empirical support for this hypothesis, which has come to be known as the "satiation effect" or "commercial wearout".

The decay hypothesis postulates that the recall decline occurs not only as the result of a change in the viewer, but also because the ad undergoes a change in its
effectiveness relative to other ads (Ostheimer, 1970). By "decay" Ostheimer means more than just "forgetting" or failing to absorb the information in the stimuli. Decay, here, basically involves a change in the ad's potency due to "extraneous" factors. Ostheimer illustrates these factors by citing a study on the effects of advertisements on physicians' drug prescriptions. Factors which decreased the effectiveness of ads were personal experience of the physicians, colleague's opinions, contacts with competitive drugs, and salesmen's calls.

Ostheimer (1970) also argues for a "reorientation of thinking" on the issue of repetition effects in advertising. The shift, he claims, must be from concern only with the total number of ad repetitions to concern for the number of repetitions within some time frame. For example, the tenth repetition in a one-hour program will quite likely have a very different effect from the tenth repetition in a one-week or one-month ad campaign. This viewpoint implies that one must address a number of variables in addition to frequency when considering repeat advertising. Among these variables are the spacing of the repetitions and the changing effects of other moderating variables during time lags. It is interesting to note that this concern is similar to what is known as the "spacing effect" or the "lag effect" in the verbal learning literature. However, it is a more global variable than the mere distribution of stimulus
presentations. It takes into account the complex and dynamic nature of advertising in a real-world setting.

Ray, Sawyer, and Strong (1971) reviewed both laboratory and field studies on the repetition effect in advertising. This review described the effects of many other variables on repeat advertising. The variables discussed include: the measure of effect or advertising goal, the audience or segment used as subjects, the product being advertised, the rest of the marketing environment, the advertising schedule. The implications of this review are clear: Any study concerned with the repetition effect in advertising must find some way to control or systematically manipulate this host of variables.

The general conclusion to be drawn from the advertising literature is that the relationship between repetition and retention is generally positive. Though the exact shape of the relationship is not clear and there is some evidence that extraneous factors interact with frequency to cause an eventual downward trend, it also appears safe to conclude that "more than a few" repetitions are probably good. The question of exactly how many repetitions are optimal has not yet been clearly answered.

Ostheimer (1970) sums up an argument favoring a large number of repetitions despite decreasing effectiveness:

If the presence of advertising has an effect, its
absence must also have an effect.... Probably the most common reaction to effect studies is that the increment produced by the greater frequency appears small. One might argue, on the other hand, that it is unrealistic to expect any increment at all, the basis of the argument being that an advertising effect at any point in time is very quickly eroded. Thus, a small increment, or none at all, becomes a reason for greater rather than less frequency. (p. 21)

Ostheimer also presents an example to illustrate his argument. Figure 1 presents the repetition function for a hypothetical ad. Presentations beyond number five only result in a small effect increment (a). However, if there are no ad presentations beyond number five, the overall effect level decreases (b). Clearly, the greater frequency is more desirable in terms of overall positive advertising effects.

Insert Figure 1 about here

On the other hand, Krugman (1972) argues that only a few exposures are optimal. In fact, he presents a speculative account (based primarily on eye camera and other "attentional research") of a viewer's reactions to ad repetitions in which only three add exposures are effective. The first exposure is a novel one which produces a response to understand the nature of the stimulus ("What is it?"). The second produces a recognition and evaluation response ("What of it?"). The third produces a response to take some sort of action (e.g., to go out and buy the product). All
Figure 1: Illustration of diminishing ad effectiveness and decay in absence of advertising.

Note: Results in Figure 1 are based upon actual data reported by Ostheimer (1970), though simplified for illustrative purposes.
other exposures are either blocked out or some "multiple" of the first three. For example, the second exposure might result in a "block out" evaluation until some time later when the 50th ad presentation becomes, subjectively, exposure number three. This formulation has not been empirically tested.

Krugman's formulation implies that only three ad exposures are psychologically effective for the viewer. The practical implications are, however, paradoxical. Since an advertiser cannot know that a viewer is "blocking out", he should continue ad repetition even though a given presentation is most likely having no effect on the viewer.

Repetition and Attitude

Because the effects of repeat advertising are often thought of in the same terms as the learning of simple verbal materials, the typical measure of effect is recognition or recall. However, Ray, Sawyer, and Strong (1971) argue that repeat advertising effects entail more than "nonsense learning", and that other criteria of repetition effects are appropriate. The most important of these other criteria are attitudes (toward the ad and toward the product) and intention to buy the product. In support of their arguments, Ray, Sawyer, and Strong conducted a series of laboratory and field studies on differential effects of ad repetition on various dependent measures. The findings of
a particular field study were typical repetition functions for ad recall measures (a negatively accelerated curve) and for brand mention or recognition (approximately linear function). However, the measure of brand preference increased only slightly over the first few exposures and then exhibited a steady negative trend. Thus, despite large positive repetition effects for retention measures, the attitude measure in this study showed no appreciable change.

The general conclusion of this series of studies was that decision-makers have only part of the picture when they do not differentiate the repetition effect for different measures and when they implicitly assume that the effect of advertising repetition is always positive. In fact, depending on the measure, the results of repetition can be positive, negative, or nonexistent. (p. 16)

Grass and Wallace (1969) report research using TV commercials and CONPAAD equipment in both laboratory and field situations. The CONPAAD device is a machine which essentially measures how much work a person is willing to perform in order to see and hear a commercial. The CONPAAD responses are considered to be indicators of the viewers attention to and interest in the ads. Grass and Wallace found that the measures of attention and retention show a typical generation/satiation cycle. However, they state that:

Attitudes, for reasons not yet understood, seem not to decline with attention but rather are maintained
near the maximum level reached at the satiation point or even increase slowly as the number of exposures increases past the satiation point. (p. 3)

In their "real world" studies, Grass and Wallace (1969) collected admittedly "crude" attitude measures and found general support for the laboratory observation that attitude measures do not satiate in the manner of retention measures. Three measures of attitude toward the product showed no decrease in favorability after the recall satiation point had been reached, although attitude toward the commercial itself did exhibit some decline. Grass and Wallace again concluded that attitude levels persist or even increase as the number of exposures exceeds the satiation point for retention. In their review article on commercial wearout, Greenberg and Suttoni (1973) make similar observations and reach similar conclusions.

In a large-scale laboratory experiment of the effects of television commercial exposure on attitude change, Winter (1973) marshals support for two hypotheses: (a) advertising exposure will have favorable effects on attitude change, and (b) each additional exposure will have a decreasing effect on that change. A more fine-grained analysis revealed that exposure effects occur primarily within the first two exposures and very little thereafter. Effects of other variables (e.g., previous exposures to ad, and brand familiarity) on attitude change were also studied.
From this study, then, it appears that attitudes toward products and commercials exhibit a negatively accelerated function of repeated exposures.

In summary, the advertising literature reveals, rather paradoxically, that while repetition functions for retention quite often follow an inverted-U curve, attitude measures increase and then asymptote. Greenberg and Suttoni (1973) offer two explanations for this paradox (i.e., that viewers' attitudes increase or maintain their level even while viewers are forgetting what they see). The first is that beyond the attention and retention satiation point viewers do not "tune out" completely from ad exposures. Greenberg and Suttoni reason that there must be some awareness in order for the viewer to decide that he has already seen the ad and to ignore it. Presumably, this fleeting exposure is enough to maintain the attitude, but not information retention. The second possible explanation is that the attitudes are somehow formed from the learned information in the ads and, once formed, may persist long after the information on which they are based has been forgotten. At best, these explanations are speculative. Possible mechanisms for the effects are not mentioned or even implied.

Ray, Sawyer, and Strong (1971) report results of one study in which no decline in attitude ratings of products
occurred during periods of no ad exposure. In fact, the ratings often became more favorable during non-exposure time. In addition, the ratings remained fairly constant after all the exposures had been completed. Their explanation of these findings is also somewhat paradoxical. They suggest that the advertising itself becomes noxious to the viewer at a certain exposure frequency. Preceding this point the effect of ad presentation on attitudes are positive, but afterwards exposures "get in the way of" favorable attitudes. In the "absence" of later ad presentations (after attention satiation), then, favorable attitudes can continue. Again, no mechanisms are proposed by which these effects might occur. It must be concluded, therefore, that explanations thus far proposed for the failure of extended repetition to reduce attitudes are insufficient and unsubstantiated.

Affective reactions to repeated stimulus exposure has been studied intensively at a basic level in addition to the above context of advertising per se. Most of the recent work stems from a monograph by Zajonc (1968) in which he presents the following hypothesis: "Mere repeated exposure of the individual to a stimulus is a sufficient condition for the enhancement of his attitude toward it" (p. 1). Zajonc defines "mere exposure" as "a condition which just makes the given stimulus accessible to the individual's perception" (p. 1). In this research, abstract stimuli are
presented repeatedly to the viewer and attitudes toward the stimuli are measured after each presentation. The general research paradigm is referred to as mere exposure research and Zajonc's hypothesis is termed the "mere exposure hypothesis" or the "frequency-affect hypothesis". Although it has only recently stimulated the interest of researchers, existence of hypotheses about mere exposure dates back to the beginnings of modern psychology (see, for example, Fechner, 1876, pp. 240-243; James, 1890).

Despite its currency, Zajonc (1968) has pointed out that direct evidence in support of the mere exposure hypothesis is lacking. He cites a number of studies concerned with music appreciation which bear some relevance, but which were the products of "antiquated methods". In addition, investigation into such complex social processes as interpersonal interaction and interracial attitudes has shown that increased exposure to other people leads one to hold more favorable attitudes toward those people (e.g. Homans, 1961; Newcomb, 1963). However, by their very complex nature, these relationships involve many intervening variables and the effects obtained cannot be directly attributed to mere exposure. Nonetheless, it is important that similar effects obtain in situations more complex than repetition of simple stimuli.

The main empirical support for the mere exposure hypothesis comes from experimental work in the area of word
frequency. Zajonc (1968) gives a very thorough documentation from his own work and that of others for the following general findings:

1. When subjects are presented with antonym pairs and asked to choose the one with the "more favorable connotation", the selected words are those used more frequently in speech and print (according to Thorndike-Lorge L-count).

2. The affective connotations of a word improve (i.e., become more positive) with repeated use.

3. Greater exposure to stimuli (words, "abstract" characters, and photographs) results in more positive attitudes toward the stimuli. In his conclusion Zajonc states that "the balance of the experimental results... is in favor of the hypothesis that mere repeated exposure of an individual to a stimulus object enhances his attitude toward it" (p. 23). Zajonc also acknowledges that further research is needed to substantiate and define the hypothesis.

Subsequent research has resulted in considerable support for the mere exposure hypothesis. However, Heingartner and Hall (1974) note that the hypothesis "has not been adequately tested using auditory stimuli, and existing results are not as clear cut as with visual stimuli" (p. 719). This contention is supported by the fact that only 15 out of 114 studies reviewed by Stang (1974) used auditory stimuli, and the majority of these suffered from antiquated
methods and other "serious methodological flaws". Consequently, Heingartner and Hall conducted a well-controlled study to test the mere exposure hypothesis with auditory stimuli. In several different settings and with two different subject populations, findings supported the hypothesis using Pakistani folk music as the auditory stimulus.

The indications are, therefore, that repeated stimuli have similar affective results whether the mode of presentation is visual or auditory. The initial findings of Zajonc (1968) showed an increasing, negatively accelerated function for the relationship between number of exposures and attitudes, or a generally linear function if the frequency axis was transformed to a log scale. Others have obtained an inverted-U function in which the intermediate frequency of exposure resulted in greatest positive affect (Stang, 1974). A discussion, from the viewpoint of social psychological theory, of why Zajonc and others may have obtained a linear function or "learning curve" instead of an inverted-U can be found in Maddi (1972, ch. 5). Basically it is argued that, for the particular stimulus employed, sufficient frequency of exposure was not used to complete the "downward" portion of the inverted-U.

In an exhaustive search of the literature on methodological factors in the frequency-affect relationship, Stang (1974) was able to evaluate statistically the effects of three man variables. First, studies which incorporated
a delay after the exposures and before the rating were more likely to obtain the enhancement of positive rating than studies which did not. In some cases, a delay of only about one minute was sufficient. In others, delays of one and two weeks were used.

Second, Stang found that the type of stimulus used played an important role. Studies which used paralogs (Turkish words, trigrams, pentagrams, etc.), ideographs, and portraits were more likely to result in enhancement than studies which used abstract visual and auditory stimuli. In addition, it was stated that stimulus complexity "seems" to be an important factor, though no solid conclusions were possible at that time.

Third, exposure paradigm was found to influence the frequency-affect relationship. That is, whether exposures are massed or distributed makes a difference. Typically, massed exposures tend to produce an inverted-U function while distributed exposures usually produce a linear "enhancement" effect. In addition, Stang considered the interaction of the exposure factor and rating delay. He suggested that no delay combined with massed exposures produces an inverted-U function, no delay combined with distributed exposures produces a flat function, and a delay with either massed or distributed exposures produces a negatively accelerated function (linearly increasing function of log
Berlyne (1970) presented sequences of colored shapes with varying frequencies and obtained ratings of "pleasingness" and "interestingness" on 7-point scales. Contrary to the mere exposure hypothesis, he found that both measures of affect decreased (i.e., became more negative) with greater stimulus frequency. In addition, when more complex stimuli were used, the opposite (that which would be predicted by the Zajonc hypothesis) occurred. Berlyne (1970) attempted to resolve these discrepancies by supposing that the changes in the hedonic value of the stimuli were dependent upon the interaction of two antagonistic factors: tedium and positive habituation. Simple stimuli emphasize tedium, resulting in the negative effects; while complex stimuli emphasize positive habituation, producing positive effects.

The two-factor theory proposed by Berlyne was extended and re-phrased by Stang (1974, 1975). According to Stang:

The theory posits that a learning and a satiation factor have additive, antagonistic roles in determining the effects of repeated exposure on affect. The theory suggests that repeated exposure is accompanied by learning about the stimulus, which in turn increases the pleasantness of the stimulus exposed; once the stimulus is learned, a boring unpleasant state of satiation is hypothesized to develop, causing a depression of affect ratings. (Stang, 1975, p. 7)

In this formulation, the different functions obtained for the frequency-affect relationship are explained by dif-
ferential predominance of the two factors. "Situations" (the methodological factors considered by Stang (1974) and any other relevant variables) which cause the learning factor to be stressed result in the typical negatively accelerated function (e.g. distributed exposures of complex, novel stimuli with delayed ratings). Satiation, which develops as learning reaches an asymptote, is stressed by large numbers of exposures and simple (i.e., quickly "learned") stimuli, and results in decreasing functions. Finally, situations which are first conducive to the learning factor and then to the satiation factor produce the inverted-U functions. The two-factor theory accounts for most of the effects summarized in Stang's (1974) review. The descriptive success lends considerable weight to the credibility of the theory, but its predictive usefullness has yet to be tested.

The basis of the two-factor theory is the link between learning and affect. The involvement of learning in the frequency-affect relationship was suggested by studies in which recognized paralogs were rated as more pleasant than unrecognized ones (Matlin, 1971) and recalled paralogs were rated more pleasant than unrecalled ones (Stang, 1973). Stang (1975) tested the relationship directly. He found that learning and affect measures have similar effects as a function of exposure frequency and duration, serial position, and rating delay. Thus, while the evidence is not conclu-
sive, it suggests that learning is a moderating variable in the relationship. This and other factors underlying the two-factor theory need further research.

It is important to note that the two-factor theory corresponds closely to the explanations of repetition effects on advertising (e.g. Grass and Wallace, 1969). In fact, some of the terminology is the same. The mediating role of learning is also an important parallel. It will be recalled that learning was found to "aid" affect (in an advertising context) up to the learning satiation point, after which the affect ratings continue to rise or remain constant while learning declines.

There are two things which seem to be consistent in the mere exposure research. One is that a generally positive relationship does, in fact, exist (at least up to the satiation point) between attitude toward a stimulus and number of exposures to the stimulus. The other is that the stimuli on which this generalization is based have, for the most part, been relatively simple. Intuitively, none of the stimuli used involve the levels of perceptual and cognitive complexity which characterize television commercials.

Research on persuasive communications utilizes more complex stimuli than those just described. Persuasive communications are complex messages which propose a main conclusion or recommendation supported by several clearly
stated arguments. Advertisements are a form of persuasive communication. The dependent variable most commonly used in this research is the subject's acceptance of the proposition or the time it takes for such agreement to occur. It is often assumed that agreement with a communication is roughly equivalent to a positive attitude rating of a stimulus. Caution is advised with this interpretation because it is not always valid to assume that acceptance indicates positive affect (Fishbein and Ajzen, 1972).

Eagly (1974) reports that the research with persuasive communications has been mostly correlational in nature, and has failed to yield consistent evidence in favor of a positive relationship between number of exposures to a message and favorable attitude change. For example, Weiss (1968) found agreement with an argument to be an increasing function of exposure frequency. On the other hand, Wilson and Miller (1968) repeated trial arguments and found a "technically significant" main effect (positive) for repetition. However, there was an interaction with the delay of the response. Apparently, there was no effect of repetition independent of this interaction. Wilson and Miller interpreted their findings as indicating that repetitive communications do not have greater persuasive impact, but that the impact decays less rapidly than that of non-repetitive communications.
Overall, evidence from social psychology substantiates the existence of a positive repetition effect for attitude measures. In mere exposure research, many of the same moderating variables and patterns of effects are found as are found in advertising research. Unfortunately, results from the research area most closely related to the advertising context, persuasive communications, are equivocal. At this time, there are indications of a positive repetition effect on attitudes, but conclusive support is lacking.

**Similarity and Attitude Repetition Effects**

It will be recalled that Grass and Wallace (1969) used the CONPAAD device to measure reactions to TV commercials. Attention, retention, and attitudes were measured as a function of repetition. Half of the experimental conditions involved straight repetition of one commercial while the other half were composed of various combinations of "similar" commercials. While the identical repetition conditions showed typical generation/satiation cycles for all three main dependent variables, the other conditions produced an interesting regeneration effect. For example, after five exposures to the same commercial, viewer interest had satiated. Introduction of a "similar" commercial on the sixth exposure, however, restored a high level of interest. Furthermore, if this commercial was "very similar" to the original commercial, a second, very
rapid, satiation occurred.

In addition to the introduction of a single similar commercial on exposure six (AAAAAB), Grass and Wallace (1969) included a condition which alternated two similar commercials (ABABAB). Compared with an overall interest loss of 46% from exposures one to five for Commercial A, the interest loss from exposures one to five in the alternation case was only 22%. When five different commercials were used (ABCDEA), the overall interest loss was only 4%.

Grass and Wallace (1969) conclude that the rate of wearout (interest decline) at any given frequency decreases as the number of different (but similar) commercials in a campaign increases. Thus, they recommend that advertisers use as many different commercials in a campaign as possible in order to prolong the delay of commercial wearout. Although it is not clearly stated, one gets the impression that recall and attitude measures followed patterns similar to those of interest measures throughout the study. In discussing the Grass and Wallace studies, Greenberg and Suttoni (1973) make similar inferences about recall and attitude measures.

It is important to recognize that "similarity" of commercials is not a completely objective concept and presents some measurement problems. In the words of Grass and Wallace:
What does or does not comprise 'differentness' would seem to be the subject of a separate investigation. But in the present study, commercials with the same objectives, theme, and approach, but based on different characters and situations, were sufficiently 'different' to decrease the rate of wearout in the simulated campaign. (1969, p. 7)

McCullough and Ostrom (1974) report a study which clearly illustrates the role of similar commercial repetition in attitude change. The authors note that previous laboratory studies have often failed to reveal positive repetition effects. However, it had been shown that people are more willing to change their attitudes when they expect a message with new information (Sears and Freedman, 1965) and that they are more attentive to a new message advocating a known position than to a previously encountered message (Brock, Albert and Becker, 1970). McCullough and Ostrom reason that the lack of a positive repetition effect in many studies is due to the repetition of identical commercials. They hypothesize that a positive effect of repetition on attitude will be obtained if the repeated commercial is "highly similar" rather than identical.

The stimuli used by McCullough and Ostrom were persuasive communications--old magazine ads which argue for the purchase of a product or solicit donations to an organization. They chose ads for each product in such a way that all employed the same general theme to support an identical position but used various topics and contained
different photographs and designs. Presentation frequency was varied from one to five and subjects' "reactions" were obtained. These reactions were later rated as favorable, unfavorable, or neutral.

Analysis of the "net cognitive response scores" supported the prediction that "attitude" is positively related to exposure frequency: i.e., similar ads produce positive repetition effects. McCullough and Ostrom (1974, p. 397) mention some important implications of these results. First, when a certain number of presentations is decided upon, it is advantageous to use as many different commercials as possible. Second, the different ads used do not have to contain different arguments, only altered contexts. These conclusions are clearly similar to those of Grass and Wallace (1969) and Greenberg and Suttoni (1973).

**Similarity and Recall Repetition Effects**

All of the applied studies above utilized attitude measures as the main dependent variable and found that ad wearout could be delayed by the use of similar ads in the repetitions. At least one of the studies (Grass and Wallace, 1969) indicated that recall measures also become more resistant to wearout as a result of similar repetitions. However, none of the studies was concerned directly with the effects of similarity on retention measures. Such effects have been investigated within the context of
verbal learning.

Martin (1968) suggests that there can be different encodings of the same nominal stimulus upon different presentations of that stimulus. The differences among the stimuli-as-coded (s-a-c) are mediated by the meaningfulness (M) of the nominal stimulus. According to Martin, variability of encoding should be greater for low-M stimuli than for high-M stimuli. In essence, a high-M stimulus is more "cohesive" in the sense that a particular meaning is nearly always attached to it. These formulations are known collectively as the encoding-variability hypothesis. It has also been suggested that varying the contextual cues associated with a nominal stimulus or altering its spatial arrangement are means of producing variable encoding.

There are two main viewpoints with respect to the effects of encoding variability on retention. One is that retention is facilitated by a stable encoding of the nominal stimulus (Martin, 1972). The basis for this position is the strength hypothesis of frequency representation which holds that the effect of stimulus repetition is to increase the strength of a single memory trace of the stimulus (Hintzman, 1969, 1970). A nominal stimulus of low variability will produce very few (ideally only one) s-a-c and each s-a-c will have a high "trace strength". The implication of this low variability effect is that the probability of a parti-
cular s-a-c being above some recall threshold is high. Thus, encoding variability should result in lower retention performance.

The other position is that retention is facilitated by variable encoding of the nominal stimulus. Repetition is effective only to the extent that it enters new information into memory on successive presentations (Madigan, 1969). Alternatively, by variable coding, repetition increases the number of access routes to the nominal stimulus (Melton, 1967; Bower, 1972). The basis for this view is the multiple-trace hypothesis which holds that the effect of stimulus repetition is to increase the number of different memory traces of the stimulus (Hintzman and Block, 1971; Howell, 1973). Each separate trace has an identity in memory which increases the probability that the nominal stimulus will be accessed for recall. Thus, encoding variability should result in higher retention performance.

Both of these contradictory predictions have received empirical support. For example, a number of investigators have found that free recall improves with the number of different contextual cues added to a repetitive stimulus (Bevan, Dukes, and Avant, 1966; Ciccone and Brelsford, Note 1; Ellis, et al., 1975). It is presumed that adding contextual cues produces variable encoding of the stimulus. On the other hand, Schwartz (1975) found that "elaborative
encoding" (nominal bigram pair associated with different word-par codes on each presentation) produced no higher recall than constant encoding (Same word-pair code on each presentation).

Several further points can be made about the equivocal findings in encoding variability research. Despite the existence of some support for Martin's hypothesis, the weight of the empirical evidence seems to favor the prediction that variable encoding enhances retention. Another point refers to the underlying representational processes for repetition. The consensus is that the multiple-trace hypothesis is a more acceptable point of view than the trace-strength hypothesis (Howell, 1973). Both of these points lend credence to the position of encoding variability effects espoused by Melton and Bower. However, for a given situation there is a trade-off with the effects of encoding variability. Variable encodings must be "similar enough" to be perceived by a subject as the same nominal stimulus in order for the retention enhancement to occur. If the encodings are so distinct as to be perceived by a subject as completely independent traces, then a strength interpretation would be appropriate to describe retention effects. This "trade-off" is based on the multiple-process hypothesis of frequency representation which holds that both strengthening and multiplexing of traces
occur concurrently, but are determined by separate factors (Howell, 1973).

The relevance of the encoding variability hypothesis for advertising repetition likes in the fact that "variability" is defined in terms of much the same operations as "similarity". A group of similar commercials usually consists of a single stimulus (e.g., the product and the main arguments) presented within different contexts (different setting, actors, etc.); thus, the series should result in multiple encodings of the primary message. The same commercial repeated a number of times should result in basically the same encoding each time. Accordingly, the variable encodings produced by similar ads should result in higher retention.

From the preceding literature review it can be seen that results from studies on stimulus repetition in an advertising context are equivocal. There seems to be a generally positive repetition effect for retention measures, but characteristics of the relationship are not well specified. When attitudes are the measure of ad effectiveness, it is not even clear that repetition effects are positive. Basic research areas in psychology which are related to advertising (verbal learning, mere exposure, persuasive communications) have shown somewhat more conclusively that repetition effects are positive. Of course, the possibility remains that these findings may not
be applicable to many complex advertising situations. In fact, when complex stimuli are used, as in persuasive communications research, definite support for a positive repetition effect is not obtained. Both basic and applied research suggest that similarity of the repeated stimuli tends to enhance positive repetition effects.

From a purely financial point of view, any contribution toward clarification of ad repetition effects would be valuable to advertisers designing a campaign. A primary consequence of possible similarity effects is that the effectiveness of a group of similar ads might be retained by altering them in repetition sequences. On the other hand, if these similarity effects are unimportant, then considerable unnecessary expense would be incurred by producing a number of different commercials instead of just one.

In addition to the practical justification, further classification of repetition effects in advertising can contribute to basic knowledge in both psychology and advertising. In particular, it would permit extension of laboratory-based hypotheses to a complex, real-world situation. The present study, therefore, was designed with the idea of examining the encoding variability and more exposure hypotheses within the context of television advertising.

Specifically, the present study was concerned with the effects of ad repetition on retention and attitude change.
The stimuli were complex, real world events—actual television commercials. They were presented to subjects within the context of a movie (carrying program) in order to simulate a common, every-day behavior—television viewing. The independent variables were number of repetitions and degree of similarity of repetitions; the primary dependent variables were recall performance and attitude change.

The primary expectations were that retention and attitude measures would increase with exposure frequency at least up to a point (satiation point), after which recall measures would decline. The effect of repeating similar (rather than identical) ads was hypothesized to enhance repetition effects either by raising the absolute level of effects, or by delaying satiation.

Method

Design

The experiment involved seven groups of subjects; six experimental groups arranged in a 2x3 between-groups factorial design and one control group. Two levels of stimulus presentation frequency (three and six) and three levels of similarity (low, medium and high) defined the groups. The control group received no repetition (i.e., a frequency of one presentation) and was included in order to provide a base level for the frequency effect.
Subjects

The subjects were 119 males and females solicited from several community and church youth groups through the groups' coordinators. Participation was on a completely voluntary basis and subjects were paid separately for the initial, primary session and for a second session (delayed posttest) at the rate of $2.00/hr.. The subjects were either paid individually or their group received a single sum on their behalf. That decision was apparently made jointly by the group members as a whole and their coordinator. With only a few exceptions the subjects were senior high school or early college students. They participated in the experiment in groups of 5 to 17, depending upon whether or not there were enough subjects for an entire experimental condition at one time. The groups were randomly assigned to experimental conditions.

Materials

The carrier program or context within which the stimuli were embedded was an old, black and white mystery movie, The 39 Steps, by Alfred Hitchcock. The stimuli were 30-second color and sound television commercials spliced into the film. The commercials were obtained from a number of advertising agencies and television stations. The target commercials advertised either Nissan Motors products (Datsun automobiles) or Rainbo bread. Buffer commercials advertised
a wide range of products and establishments. To control for the effects of previous exposures to the target commercials, they were chosen to satisfy the following criteria: (a) the products advertised were nationally distributed, (b) the particular commercials were part of a national campaign, and (c) the individual commercials were, preferably, at least six years old. If the last condition could not be met, as was the case in a few instances, then the commercials were brand new and not yet widely distributed on television. The buffer commercials did not necessarily meet these criteria.

The target commercials were all supportive in nature. The different levels of similarity were operationally defined as follows:

1. High—identical commercial repeated.
2. Medium—"highly similar" commercials repeated (same product, same claims put forth about product, same general approach, but slightly different format).
3. Low—"similar" commercials repeated (same product, different approach, completely different format, different attributes of product stressed).

In the case of Datsun automobiles, enough appropriate commercials for one model could not be obtained to meet the criteria for low similarity. Instead, commercials for different automobile models were used, thus meeting the
criteria for the brand (Datsun) as opposed to the particular product.

The subjects' responses were collected in four standardized paper and pencil "response booklets" (see Appendix A for sample Response Booklet). Response Booklet 1 was a pretest which contained questions pertaining to general demographic data and television viewing preferences. In addition, it contained scales for rating prior attitudes toward the target products and a number of buffer products. Attitude ratings were obtained on six 7-point semantic differential scales and two 7-point scales designed to measure overall attitude toward the specific commercials and the subject's intention to buy the products. Four of the semantic differential scales were chosen to have high loadings on the Evaluative factor, while the other two had high loadings on the Activity and Potency factors.

Response Booklet 2 was a posttest which consisted of questions pertaining to recall of the movie, information in target and buffer commercials, and judgment of frequency. The commercial recall questions were arranged in such an order as to insure that cued recall questions followed free recall questions. In addition, blank pages were built into the booklet as shields and subjects were specifically instructed not to work any of the pages out of sequence.

Response Booklet 3 was a posttest which consisted of
attitude ratings of target and buffer commercials as in Booklet 1 above. Response Booklet 4 was a delayed posttest which consisted of a combination of Booklets 2 and 3 in shortened form (questions and ratings pertaining to the movie and buffer commercials were deleted).

For all the attitude-rating responses, it was decided that only the four evaluative scales (i.e., those semantic differential items that loaded heavily on the Evaluative dimension) and the commercial-rating item would be utilized. This decision was based primarily upon the results of pilot research which was not completely analyzed at the time the present study was undertaken. Further, it was dictated by the fact that subjects had considerable difficulty rating various products on the three items which were not predominantly evaluative in nature.

The questions concerned with recall (Booklets 3 and 4) were of four types, operationally defined as follows:

1. Free Recall--straight free recall of all commercials seen (brand name and product).

2. Cued Recall--subject given cue of generic product type ("automobile" or "bread product") and asked to recall brand name and product.

3. Free Recall, Information--subject given cue of specific commercial ("You saw a Datsun automobile commercial") and asked to recall as much specific information as
possible.

4. Cued Recall, Information—subject given cues of specific commercial and questions about specific information contained therein ("Answer the following questions about the Datsun automobile commercial(s) you saw... ").

After construction of the booklets, standardized scoring rules were devised for scoring the responses. In the case of attitude ratings this simply involved attaching an appropriate number to each segment of the scale. In all cases "1" corresponded to the positive end of the scale and "7" to the negative end. In the case of recall questions, the following criteria were used to score each type of question:

1. Free Recall--one point was given for correct identification of the brand and product for each target product (e.g. "Rainbo bread" and "Datsun sedan" = 4 points).

2. Cued Recall--same as 1. above, except the particular product was required.

3. Free Recall, Information--one point given for each individual item of information recalled by the subject (varied according to commercial(s) used in each condition).

4. Cued Recall, Information--one point given for each question correctly answered by subject (varied according to commercial(s) used in each condition).
Procedure

On each session, the subjects were seated in a lecture hall and told that they were participating in a study designed to investigate the effects of television viewing. They were informed that they would see a mystery movie which would be interrupted at various times by commercials—a simulated television viewing situation. The subjects were instructed to complete Response Booklet 1, following which the film was shown.

There were seven designated commercial "spots"; one before, one after, and five during the movie. The choice of spots during the movie was predetermined in order to coincide with a logical break in the movie (e.g. a scene change) and also to occur at approximately equally spaced intervals. During each spot, four 30-second commercials were shown. After the seventh spot, Response Booklets 2 and 3 were administered immediately and in succession. After completing Booklet 3, subjects were informed that a final booklet would be distributed one week later. Approximately one week later, Booklet 4 was administered to as many subjects as were available. All subjects were paid upon completion of the first session; those who participated in the delayed session were paid an additional amount at the close of that session.

The commercial spots were constructed so as to maintain, as much as possible, constant spacing for the target commerc-
cials. All of the target commercials were "spaced" as opposed to "massed". That is, no one spot contained more than a single target commercial for the same product. Within a spot, order of commercials was randomly determined except that target commercials were always the second and third items (in order to insulate the target commercials from immediate movie context effects). The locations of spots in which target commercials appeared were as follows: For a frequency of one, target commercials appeared in the fourth spot. For a frequency of three, target commercials appeared in the second, fourth, and sixth spots. For a frequency of six, target commercials appeared in the second through the seventh spots. The buffer commercials remained constant, except where target commercials were inserted for different conditions. See Appendix B for a sample commercial line-up.

Results

The study was designed as a 2x3 factorial (frequency x similarity) with a single control group (frequency of one) included in order to provide baseline data for the repetition effect. The single control group prevented frequency from being fully crossed with similarity. However, Winer (1962, pp. 264-267) discusses a method for including such a control group in a complete factorial design. Basically, he suggests using the control group data several times in
order to create artificial crossing with the other factors. Following this procedure, the data from the present control group were included at all three similarity levels, creating (artificially) a third level of frequency. Data were therefore analyzed, in most cases, using a $3 \times 3$ factorial design. In order to compensate partially for the resulting lack of statistical independence among the replicated cells, a modified error term and df calculation were used. This was accomplished by casting the original data into the form of a one-way ANOVA (all treatments combined and computing the error term; then subtracting one df from that error estimate for each replicated cell in the $3 \times 3$ design.

Himmelfarb (1975) suggests another solution in which the control group is subdivided and randomly assigned to each additional cell of the complete factorial design. This method was not used because of an insufficient number of subjects in the control group.

Recall Data

The recall performance was measured in terms of four dependent variables: free recall, cued recall, free recall-information, and cued recall-information. The cued recall results were not analyzed because it was discovered that this measure was confounded with the amount of information in the stimuli in several conditions. Because the recall results were very similar for both target products, the
data were summed for the two products on each of the recall variables.

--- Insert Figure 2 about here ---

The number of items recalled for the free recall measure is presented as a function of frequency and similarity in Figure 2. It can be seen that recall improves with increased frequency, but apparently is little affected by similarity. An ANOVA supported these conclusions. The main effect for frequency was significant, $F(2,110)=21.76$, $p<.001$, while the main effect for similarity was not, $F(2,110)>1$. The interaction of frequency with similarity did not approach significance, $F(4,110)>1$.

--- Insert Figure 3 about here ---

The mean number of items recalled for the free recall-information measure is presented in Figure 3. It can be seen, again, that recall improves considerably with increased frequency. However, it appears that high similarity produces higher recall than either low or medium similarity, and that the similarity effect interacts with frequency. These conclusions were borne out by an ANOVA: The main effects of frequency, $F(2,110)=29.032$, $p<.001$, and similarly, $F(2,110)=5.178$, $p<.01$, were significant; as was the interaction between frequency and similarity,
Figure 2. Mean Free Recall performance as a function of frequency and similarity.
Figure 3. Mean Free Recall-Information performance as a function of frequency and similarity.
The percentage of items correctly recalled for the cued recall-information measure is presented in Figure 4. Once again it appears that recall bears a strong positive relationship to frequency. It also appears that similarity affects the recall performance and that this effect is mediated by frequency. These conclusions were substantiated by the ANOVA which showed significant effects for frequency, $F(2,110)=97.68, p<.001$; similarity, $F(2,110)=5.95, p<.01$; and the interaction, $F(4,110)=2.46, p<.05$.

Planned comparisons were carried out when there was a significant main effect for frequency in order to determine if the effect continued to be significant between the frequencies of three and six. The purpose of this was to reveal satiation effects (curve asymptotes). Standard F-ratios were constructed using the same corrected error MS and df as above. The individual comparisons are presented in Table 1. It can be seen that every recall measure resulted in a significant increase between frequencies three and six for both low and medium similarities, but no significant changes between these frequencies in any of the high similarity conditions.
Figure 4. Mean Cued Recall-Information performance as a function of frequency and similarity.
Table 1
Planned Comparisons Between Frequencies of Three and Six for Measures Having Significant Main Effect of Frequency

<table>
<thead>
<tr>
<th>Dependent Measures</th>
<th>Similarity</th>
<th>F-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free Recall</td>
<td>Low</td>
<td>7.602*</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>5.505**</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>0.155</td>
</tr>
<tr>
<td>Free Recall</td>
<td>Low</td>
<td>4.217**</td>
</tr>
<tr>
<td>Information</td>
<td>Medium</td>
<td>11.040*</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>0.201</td>
</tr>
<tr>
<td>Cued Recall</td>
<td>Low</td>
<td>19.180*</td>
</tr>
<tr>
<td>Information</td>
<td>Medium</td>
<td>6.960**</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>2.630</td>
</tr>
<tr>
<td>Rainbo Commercial</td>
<td>Low</td>
<td>1.544</td>
</tr>
<tr>
<td>Ratings</td>
<td>Medium</td>
<td>17.054*</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>4.263**</td>
</tr>
</tbody>
</table>

\(^{a}df\ (1,32)\)

* \(p < .01\)

** \(p < .05\)
Attitude Data

Attitude ratings were taken for both the advertised target products and the target ads. The product ratings were summed for each subject across the four evaluative scales. Change scores (pretest-posttest) were calculated for both measures and subjected to an ANOVA. Because the results differed considerably for the two target products, each was analyzed separately.

The mean attitude change scores are summarized in Table 2. In the case of Datsun products, there appear to be no marked differences attributable to the experimental treatments. In fact, an ANOVA of these data yielded no $F$ values which even approached significance: For frequency, $F(2,110) > 1$; for similarity, $F(2,110) > 1$; and for interaction, $F(4,110) = 1.334$, $p > .25$. Since most of the change scores were positive, it is possible that there was a slight positive shift in attitude as a result of ad exposure regardless of frequency. However, a test of the combined effects revealed that the shift was not significantly different from zero, $t(118) = 1.33$, $p > .05$.

The ratings of Datsun commercials exhibit results similar to those above. An ANOVA of these change scores yielded a frequency effect which approached significance,
Table 2

Mean Attitude Change Scores for Different Products and Commercials as a Function of Frequency and Similarity

| PRODUCT SIMILARITY | FREQUENCY | | |
|---------------------|-----------|---|---|---|
|                     | 1         | 3 | 6 | |
| PRODUCT RATINGS     |           |   |   |   |
| Low                 |           |   |   |   |
| Datsun              |           |   |   |   |
| Low                 | +0.65     | +0.41 | +0.41 |
| Medium              | +0.65     | 0.00 | +0.82 |
| High                | +0.65     | +1.65 | -1.12 |
| Rainbo              |           |   |   |   |
| Low                 | +5.18     | +0.41 | -1.29 |
| Medium              | +5.18     | +2.00 | +3.24 |
| High                | +5.18     | -2.29 | +2.06 |
| COMMERCIAL RATINGS  |           |   |   |   |
| Datsun              |           |   |   |   |
| Low                 | -0.06     | +0.82 | +0.29 |
| Medium              | -0.06     | +0.65 | +0.41 |
| High                | -0.06     | +0.41 | -0.59 |
| Rainbo              |           |   |   |   |
| Low                 | +1.29     | +1.18 | +0.47 |
| Medium              | +1.29     | +1.59 | -0.77 |
| High                | +1.29     | +0.94 | -0.24 |
\( F(2,110)=3.05, p<.06 \), but a similarity effect and interaction which were well below significance \( F(2,110)=1.315, p>.25 \) and \( F(4,110)>1 \), respectively. In this case, however, a test of the combined effects of ad exposure revealed that attitude change is positive and significantly different from zero, \( t(118)=1.98, p=.05 \).

In the case of Rainbo products there appears to be a definite trend toward reduced commercial effectiveness as a function of frequency. That is, changes in attitude as a result of exposure to the advertising become less positive, even negative in some cases, as frequency is increased. The ANOVA revealed that this frequency effect was indeed significant, \( F(2,110)=15.49, p<.001 \). Although the groups also seem to differ as a function of similarity, the differences are rather unsystematic. For example, the medium and high similarity groups exhibit a slight negative shift with frequency; the low similarity group, a large negative shift. This similarity effect was reflected in the ANOVA, which approached significance, \( F(2,110)=2.79, p<.07 \). Similarly, the interaction effect, while weak, also approached significance, \( F(4,110)=2.25, p<.07 \). The overall positive effect (for all the groups combined) did achieve significance, \( t(118)=2.73, p<.01 \).

The Rainbo commercial ratings show the most marked effects of all the attitude measures. Once again, the trend
is for the positive change in attitude ratings to decrease in magnitude (and become negative) as a function of frequency. This effect was found to be significant, $F(2,110)=3.93$, $p<.05$. It is also apparent that the attitude ratings vary as a function of similarity. For example, medium and high similarity groups show a greater decline than the low similarity group. This effect was also found to be significant, $F(2,110)=6.887$, $p<.01$. There appears to be an interaction between frequency and similarity. In particular, the change scores for a frequency of three are nearly equal across similarity, but for a frequency of six they differ considerably along similarity levels. These observations were supported by a significant interaction effect from the ANOVA, $F(4,110)=2.465$, $p=.05$. Though the general trend is one of decreasing commercial effectiveness, the entire mean rating change was positive and highly significant, $t(118)=3.93$, $p<.001$.

The only planned comparisons carried out with the attitude data were between frequencies three and six for the Rainbo commercial ratings. Table 1 shows that the rating decline stops in the low similarity conditions, but continues in the other two similarity conditions.

Discussion

The present results clearly suggest that advertising effect, when measured by recall, is enhanced through
repetition. All three measures of recall exhibited this trend across all levels of similarity with the differences being highly significant. This general finding supports the expectation derived from both basic (e.g. verbal learning) and applied (e.g. advertising) research settings that retention increases as a function of frequency.

When attitude scales are used to measure the effect of advertising, the repetition function is not as clear-cut as with retention measures. To a degree, it appears to depend upon the specific context of the advertising. In the case of attitudes toward the products, for example, repetition of Datsun ads produced no significant changes in subjects' ratings in any of the analyses performed. On the other hand, repetition of Rainbo ads produced two major effects. First, an overall test showed that mere exposure of the ad, independent of frequency, had a positive effect on attitude. Second, an ANOVA revealed a negative relationship between attitude and frequency. That is, initially positive attitude ratings declined as a function of frequency.

In the case of attitudes toward the commercials the repetition effect was similar to that for the products, but somewhat more pronounced. The relationship between Datsun ad ratings and frequency was positive, but somewhat below the level necessary for statistical significance ($p=.05$). A single test of all change scores combined
revealed, once again, that the overall shift was significantly greater than zero in the positive direction. In other words, ad presentation in general had a positive effect on attitudes, but this effect was not reliably enhanced by greater frequency. The ratings for the Rainbo ads were very similar to the Rainbo product ratings. The combined change scores were significantly greater than zero, but the positive attitude change decreased significantly as a function of frequency.

While retention measures thus show a clear, positive repetition effect, the attitude measures fail to confirm this finding. In fact, the present results suggest that the relationship between attitudes and ad repetition can even be negative under some circumstances. This is apparently in direct contradiction to the large number of studies on mere exposure and advertising effectiveness which have shown a positive frequency-affect relationship. However, all previous research has not supported the positive relationship. The basic research on persuasive communication, for example, has not produced positive effects consistently. More importantly, at least one study conducted within an advertising context reported both negative and zero repetition effects (Ray, Sawyer, and Strong, 1971). Since television commercials are a form of persuasive communication, it is not surprising that atti-
tudes failed to become progressively more favorable with frequency in the present study.

An important inference can be drawn from the two separate analyses of the attitude data. The ANOVA revealed negative or non-existent frequency effects, while the combined change score tests revealed overall positive frequency effects. Taken together, these analyses suggest that the initial effect of ad exposure on attitude may be positive, but that these positive effects are not enhanced by--or may even decrease with--subsequent exposures. Interpreted in this manner, the present data provide an indication of a satiation in effectiveness with respect to attitudes. This finding is typical of previous advertising recall research, but not attitude data. Further, if, in fact, the initial attitudes were positive and no enhancement occurred as a result of ad repetitions, then satiation began almost immediately. This finding supports those by Krugman (1972) and Winter (1973) that ad effectiveness spans only about three exposures.

Although several aspects of the present attitude data can be interpreted as support for hypotheses such as Krugman's and the satiation effect caution should be used in making such an interpretation. It should be recognized that the present attitude results are not very conclusive overall. A number of factors may have contri-
buted to the inconclusive nature of the findings. First, there is the general problem of attitude measurement. The semantic differential scales have demonstrable validity for measuring dimensions of connotative meaning—especially evaluation. However, it cannot be assumed that such items measure affect equally well under all circumstances; items may take on somewhat different meanings for different concepts and different subjects (Fishbein and Ayzen, 1972).

A related factor which may have contributed to the attitude results is the validity of the Evaluative dimension of the semantic differential for application to advertisements. There has been a vast amount of research on the construction and validation of semantic differential scales. It has been demonstrated that the evaluative items reliably measure a coherent evaluative aspect of meaning. It seemed reasonable, then, to sum across a number of items that load almost exclusively on the Evaluative factor to obtain a measure of attitude toward any concept. However, some recent evidence suggests that such a procedure may not be completely satisfactory applied to television ads.

A factor analytic study of Leavitt (1970) uncovered seven primary dimensions on which such ads could be rated. No single evaluative factor appeared. Rather, any affective component was distributed over the seven factors. This implies that the Evaluative factor might not be a unitary
dimension for television ads. Consequently, it is possible that each of the four evaluative items used in the present study was somewhat independent. If this were the case, then summing all four scales could have obscured systematic effects.

A third explanation for the monosystematic attitude results is the potential confounding effect of previous exposure. While it is true that the study controlled to a degree previous exposure to the specific ads used as stimuli, it was impossible to control for more "generic" previous exposure (e.g. other car or bread company ads) or individual viewing history. Although situational, rather than background frequency (the analog of previous ad exposure) is the dominant factor in verbal learning research, background frequency may have more of an effect on attitudes in an advertising context. Surely the subjects had been exposed previously to related ads. It seems reasonable, then, that the subjects could have been "on the verge of" satiation prior to the present ad exposures. The fact that some of the data could be interpreted as displaying satiation very quickly would lend some credence to this possibility. Obviously, an effect such as this would help explain the present absence of conclusive frequency effects. Further research into this matter would certainly be in order.
It will be recalled that a primary expectation was that lower similarity repetition would tend to enhance the frequency effects. Two means were hypothesized by which this effect could occur: either the lower similarity levels would result in greater ad effectiveness (i.e., higher recall or more positive attitudes), or the lower similarity levels would result in delayed or eliminated satiation. Because the attitude data were equivocal with respect to identical repetition effects, no firm conclusions could be drawn concerning this similarity effect. The one instance which exhibited a significant similarity effect (Rainbo commercial ratings), however, lends tentative support to this expectation. All three similarity levels showed a decline in positive attitude ratings, but the low similarity group exhibited a more gradual decline than the other two groups.

The recall data, on the other hand, provide some support for the hypothesized similarity effect. In general, the expectation that lower similarity would result in better recall was not borne out. Inspection of Figures 2 through 4 reveals that high similarity produces a higher absolute level of recall than medium or low similarity (the latter two being very comparable). However, in terms of the other criterion of enhancement (lower similarity delaying satiation), the data do support the similarity hypothesis.
None of the three high similarity functions showed a significant change between frequencies three and six. Apparently, the curves had asymptoted by this time and satiation had begun. On the other hand, recall for the low and medium similarity levels increased significantly between frequencies three and six in every case—the curves showed no indication of satiating or reaching an asymptote.

These two points taken together imply that the real benefit of lower similarity repetition is realized at frequencies greater than six. It appears that the low and medium similarity curves would have crossed over and exceeded the level of the high similarity curve had the frequency been extended. This trend is apparent in all three recall measures and the crossover even seems to have occurred already in the case of the free recall measure.

The importance of the similarity effect extends to basic theoretical notions. It provides important empirical support in a complex, real-world setting for the encoding variability hypothesis. All that is necessary to interpret this study in terms of encoding variability is to "translate inversely" from similarity to variability (e.g. high similarity is equivalent to low variability). Apparently, more variable encoding (induced by context) results in the enhancement of recall. Again, the expected increase in terms of absolute recall level did not occur, but trends
in that direction were apparent. This observation also bears on the issue of which interpretation of encoding variability effects is more plausible: The initial data analysis suggests that Martin's strength approach is correct in predicting better recall with low variability, but further analysis suggests that the opposite (multiplexing interpretation by Melton and Bower) would be true if frequency were extended. It would seem more appropriate, from the standpoint of the present data, to invoke a multiple-process viewpoint of encoding variability which posits a trade-off between trace strength and multiple trace accumulation. In this study, it seems that trace strength built up rapidly for the low variability condition, while it required more time (i.e., greater number of exposures) for multiple traces to accrue (and to dominate the trade-off) in medium and high variability conditions.

In summary, the present data provide support for the expectation that retention increases as a function of ad repetition. But the data do not substantiate a similar expectation for attitude change, thereby failing to support the mere exposure hypothesis as applied to repeat advertising. Further, as expected, repetition of identical ads resulted in a satiation or wearout effect--the recall-frequency function approached asymptote after a relatively small number of exposures. As hypothesized, the repetition
of similar, rather than identical, ads ameliorated the satiation effect. Thus, support was provided for an encoding variability hypothesis as applied to repeat advertising.

One practical implication of these findings is that advertisers should use discretion before going through the expense of producing several similar ads instead of repeating a single ad. If an advertiser wishes viewers to remember an ad and its contents and a campaign is designed to cover a short time period (i.e., few exposures), then the present results would indicate that identical repetition would be more effective than non-identical repetition. If the campaign is to be a longer one, then non-identical repetitions would be more effective.

The present study could be replicated and improved in a number of ways. The use of other techniques for attitude measuring might shed light on the issue of whether or not the evaluative rating scales are useful in an advertising context. For example, comparisons could be made with data from the pupillary response. Greater attention must be paid to what dimensions of advertisements are measured by rating scales and other devices. In fact, more precise knowledge about the actual components of a stimulus as complex as an ad is not only needed for more accurate measurement of dependent variables, but also for determining levels
of independent variables. In this study, similarity was scaled on the basis of previous research and intuition. The anticipated effects of non-identical repetition were achieved, but the two levels of non-identical repetition (low and medium similarity) were not differentiated. In other words, it is necessary to scale degrees of ad similarity *empirically*. The effects studied in the present contexts need to be studied at higher frequency levels. There were clear indications that important trends might have developed at higher frequencies. An attempt was made to control for the effects of previous exposure by using only older commercials which were nationally distributed. An alternative to this method would be to use all brand new ads. Finally, this study should be replicated using various other advertisements and manipulating other variables together with frequency and similarity.
Reference Note

References


Himmelfarb, S. What do you do when the control group doesn't fit into the factorial design? *Psychological Bulletin*, 1975, 82, 363-368.

Hintzman, D. L. Apparent frequency as a function of frequency and the spacing of repetitions. *Journal of Experimental Psychology*, 1969, 80, 139-145.


Wilson, W. and Miller, H. Repetition, order of presentation, and timing of arguments and measures as determinants of opinion change. *Journal of Personality and Social Psychology*, 1968, 9, 184-188.


Appendix A

Sample Response Booklet: Booklet 4 -- Groups 6 and 7
similarity=medium
frequency=3 and 6

Page 1:
NAME_____________________________________

INSTRUCTIONS:

On each page of this booklet there are questions concerned with the simulated television programming you watched. Please try to answer all of them as best you can. On questions which require that you write in answers, please do so as legibly as possible. Correct grammar and spelling are not important, but be very specific in your answers. On questions which require that you rate some product on various scales, do so as before, relying on your first impression.

Do each page in order. It is absolutely essential that you do not look ahead or back to other pages.

Page 2:
Write in the names of as many brand names and particular products as you can remember that appeared in the commercials you saw. If you can only remember the brand name or the product, write in that much (e.g., Starkist Tuna, Starkist____, or ____ Tuna.). If you need more room, use the back of the preceding page

Page 3:
Write down the name of any automobiles you saw a commercial for. Indicate the company and the particular model.

Page 4:
Write down the name of any bread product you saw a commercial for. Indicate the company and the particular product.
Page 5:
Write down as much specific information as you can remember about the Datsun automobile commercial(s) you saw.

Page 6:
Write down as much specific information as you can remember about the Rainbo bread products commercial you saw.

Page 7:
Answer the following questions about the Datsun automobile commercial you saw:

1) What general claims are made about the Datsun?
2) What was the EPA mileage rating of the Datsun?
3) What was the model number of the advertised Datsun?
4) What is the company's motto?

Page 8:
Answer the following questions about the Rainbo bread product commercial(s) you saw:

1) What is the motto for Rainbo bread as it appears on the package?
2) How long do the described recipes take to prepare?
3) How long is the baking process for Rainbo bread (to set it apart from other breads)?
4) How long do other bread companies take to make their bread?
Page 9:

In the space following each item below, indicate in numbers how many times a commercial for that particular product was shown:

Ban Roll On  ______  Crisco Shortening  ______
Jergens Hand Lotion  ______  GE Products  ______
Lone Star Beer  ______  Texaco Gasoline  ______
La Choy Chow Mein  ______  Kraft BBQ Sauce  ______
Commet Cleanser  ______  Datsun Autos  ______
Serta Mattresses  ______  Rainbo Bread  ______

Page 10:

DATSUN AUTOS  Good:____:____:____:____:____:____:____:Bad
Honest:____:____:____:____:____:____:____:Dishonest
Valuable:____:____:____:____:____:____:____:Worthless
Pleasant:____:____:____:____:____:____:____:Unpleasant
Young:____:____:____:____:____:____:____:Old
Strong:____:____:____:____:____:____:____:Weak
Good Commercial:____:____:____:____:____:____:____:Bad Commercial
Likely To Buy:____:____:____:____:____:____:____:Unlikely To Buy
Page 11:

RAINBO BREAD

Good: __:__:__:__:__:__:__:__:__:__:Bad
Honest: __:__:__:__:__:__:__:__:__:__:Dishonest
Valuable: __:__:__:__:__:__:__:__:__:__:Worthless
Pleasant: __:__:__:__:__:__:__:__:__:__:Unpleasant
Young: __:__:__:__:__:__:__:__:__:__:Old
Strong: __:__:__:__:__:__:__:__:__:__:Weak
Good Commercial: __:__:__:__:__:__:__:__:__:__:Bad Commercial
Likely To Buy: __:__:__:__:__:__:__:__:__:__:Unlikely To Buy
Appendix B

Group 4, 7/31/75
frequency=3, similarity=high

<table>
<thead>
<tr>
<th>SPOT NUMBER</th>
<th>ORDER OF COMMERCIALS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Kimberly Clark #5</td>
</tr>
<tr>
<td></td>
<td>Kraft BBQ Sauce #29a</td>
</tr>
<tr>
<td></td>
<td>University Savings #53a</td>
</tr>
<tr>
<td></td>
<td>La Choy #2</td>
</tr>
<tr>
<td>2</td>
<td>G.E. Dishwasher #52</td>
</tr>
<tr>
<td></td>
<td>Datsun #62a</td>
</tr>
<tr>
<td></td>
<td>Rainbo #69a</td>
</tr>
<tr>
<td></td>
<td>Crisco Oil #1 M.C.</td>
</tr>
<tr>
<td>3</td>
<td>Comet #2 M.C.</td>
</tr>
<tr>
<td></td>
<td>Ajax #11 M.C.</td>
</tr>
<tr>
<td></td>
<td>Brightside #12 M.C.</td>
</tr>
<tr>
<td></td>
<td>Liquid Drano #5 M.C.</td>
</tr>
<tr>
<td>4</td>
<td>Houston Library #6 M.C.</td>
</tr>
<tr>
<td></td>
<td>Datsun #62b</td>
</tr>
<tr>
<td></td>
<td>Rainbo #69b</td>
</tr>
<tr>
<td></td>
<td>Bufferin #7 M.C.</td>
</tr>
<tr>
<td>5</td>
<td>Philco-Ford (1 min) #1</td>
</tr>
<tr>
<td></td>
<td>Kraft Pizza #30a</td>
</tr>
<tr>
<td></td>
<td>Fab #32 M.C.</td>
</tr>
<tr>
<td>6</td>
<td>Kraft Said Dressing #34</td>
</tr>
<tr>
<td></td>
<td>Rainbo #69c</td>
</tr>
<tr>
<td></td>
<td>Datsun #62c</td>
</tr>
<tr>
<td></td>
<td>Stretch N' Seal #45 M.C.</td>
</tr>
<tr>
<td>7</td>
<td>Lone Star Beer #47 M.C.</td>
</tr>
<tr>
<td></td>
<td>Vanish #50 M.C.</td>
</tr>
<tr>
<td></td>
<td>Nice N' Easy #52 M.C.</td>
</tr>
<tr>
<td></td>
<td>Ban Roll On #53 M.C.</td>
</tr>
</tbody>
</table>

NOTE: Target commercials underlined.