**A Meta-Analysis of the Spacing Effect in Verbal Learning:
Implications for Research on Advertising Repetition and Consumer Memory**

CHRIS JANISZEWSKI*
HAYDEN NOEL
ALAN G. SAWYER

May, 2002

* Chris Janiszewski is the Jack Faricy professor of marketing, University of Florida, Gainesville, FL 32611-7155 (chrisj@dale.cba.ufl.edu). Hayden Noel is doctoral student, University of Florida, Gainesville, FL 32611-7155 (noelhn@chip.cba.ufl.edu). Alan Sawyer is J.C. Penney professor of marketing, University of Florida, Gainesville, FL 32611-7155 (asawyer@dale.cba.ufl.edu).

The effects of repeated advertising exposures depend on the size of the interval, or space, between ad exposures. A meta-analysis of 97 verbal learning studies identified several stimulus characteristics and learning context factors that interact with stimulus spacing to facilitate memory for repeated information. The majority of the findings are consistent with the predictions of two enhanced processing explanations of learning – the retrieval hypothesis and the reconstruction hypothesis. These two hypotheses predict that an effective repetition strategy should encourage incidental processing during one presentation of the material and intentional processing during the other presentation of the material, but the hypotheses differ about the optimal order of these two types of processing. Thus, the most effective repetition strategy may be a combination of spaced exposures that alternate in terms of media that is involving (e.g., television commercials) and less involving (e.g., billboards, product placements).

Repetition is a fundamental advertising strategy that can be used to achieve several goals (Pechmann and Stewart 1989; Ray and Sawyer 1971a; Unnava and Burnkrant 1991). Repetition increases the likelihood a target market will attend to an ad, remember its content, and be persuaded by its message. At the most fundamental level, increased memory for a brand name or product benefit results in an increased probability that the brand will enter into the consumer's consideration set.

Recommendations about how to make advertising more memorable often focus on characteristics of the ad material, competing ads, and the audience (Sawyer 1981). What is often neglected is the potential for an interaction between the advertising content or context and repetition. For example, memory for simpler, more novel, or more meaningful advertising material may be enhanced more by an initial exposure, whereas memory for more complex, more familiar, or more abstract advertising material may be enhanced more by additional exposures. An understanding of how different types of advertising are learned and remembered as a consequence of repeated exposure can provide insight into how to make these communications more effective learning tools**.**

One important aspect of advertising campaigns that use repetition is the scheduling of the repeated exposures. Formal research about how repetition schedules interact with advertising content and context is sparse. Advertising researchers rarely manipulate advertising content and context in conjunction with different repetition schedules, be it within an ad or across flights of ads (for some exceptions, see Rethans, Swasy, and Marks 1986; Schumann, Petty, and Clemons 1990; Ray and Sawyer 1971a). Even when content and context variables are investigated, the memory effects associated with different levels of these variables are often tangential to the main

focus of the study (e.g., Rethans et al. 1986; Schumann et al. 1990). As a consequence, conclusions about repetition and memory tend to be generalizations about characteristics of the repetition schedule. For example, more repetition is better than less repetition; highly concentrated repetition may be damaging; and distributed presentations work better than massed presentations (Craig, Sternthal, and Leavitt 1976; Malaviya and Sternthal 1997; Ray and Sawyer 1971a; Zielske 1959). Thus, the literature provides insufficient insight as to why memory for some ads benefits more from repeated exposure than memory for others.

An important characteristic of ad scheduling is the duration of time or number of competing ads separating one exposure from another. Given the dearth of advertising research about this topic, a strategy for gleaning insights about how different exposure intervals interact with changes in advertising content and context is to investigate literatures that actively manipulate these variables in tandem. For example, the verbal learning literature has over 40 years of documented investigations into learning via repetition. More specifically, the spacing literature has investigated the influence of increasing the time interval between two presentations of a stimulus on memory for the stimulus. Although this literature lacks the external validity of field studies investigating ad campaigns, it has the internal validity needed to make inferences about what forms of ad content, and in what contexts, should most benefit from repeated exposure.

Research on the spacing effect compares a *massed* condition, where repeated items are presented contiguously, to a *distributed or spaced* condition, where repeated items are separated by intervening material, tasks, or time. In general, a distributed presentation schedule results in better memory for the items than a massed presentation schedule, a result commonly called the *spacing effect*. Moreover, as the interval between the repeated presentations of a stimulus

increases, memory for the item increases at a decreasing rate, a result commonly called the *lag effect* (Ray and Sawyer 1971b). The spacing of stimulus presentations has been shown to enhance memory for nonsense syllables (e.g., Ebbinghaus 1885), words (e.g., Glenberg and Lehmann 1980), sentences (e.g., Rothkoph and Coke 1966), pictures (e.g., Hintzman and Rogers 1973), faces (Cornell 1980), and advertisements (Zielske 1959).

We expect that a meta-analysis of the spacing effect will contribute to the advertising and consumer behavior literature in two ways. First, it will provide insights into relationships between ad content or context, ad repetition, and memory. For example, we know that manufactures can benefit from selecting a meaningful word as a brand name because of the associations it evokes (Keller, Heckler, and Houston 1998). We also know that Proctor and Gamble often uses meaningless words as brand names and that many global competitors have meaningless brand names to many potential consumers in a new market. Thus, we have an opportunity to disentangle the initial memory benefits of meaningful words from the potential memory benefits of repeated exposure to different classes of words. We also know that varied ad executions can discourage a drop in recall scores (Greenberg and Suttoni 1973), but we know less about how to vary ad executions to promote memory, or if in fact, varied ad executions will promote memory.

The second goal of the meta-analysis is to provide insight into how people learn via repetition. In the marketing literature, the most popular explanation of learning via repetition is encoding variability theory (cf. Singh et al. 1994; Unnava and Burnkrant 1991). Encoding variability theory predicts that presenting a series of ads consisting of slight variations on a theme (e.g., the Absolute vodka ad campaign) enhances memory for the ad material. Although the results of the limited research in the marketing literature have been consistent with the encoding

variability hypothesis, there has been mounting empirical support for alternative hypotheses in the spacing literature (Dempster 1996; Postman and Knecht 1983). Identifying competing theories that can account for the empirical evidence related to advertising repetition may provide further insight into the memory processes responsible for learning.

This article is organized as follows. First, we discuss five of the explanations that have been proposed to account for the spacing effect. Next, we offer ten hypotheses implied by one or more of these theories, paying special attention to stimulus and context features that differentiate among the explanations. Then we present the results: 97 independent studies provide 269 data points for testing the hypotheses. We identify relationships between processing goals, stimulus characteristics, and the spacing of stimulus presentations and discuss how these relationships may provide insight into how consumers form memories of marketing information.

SPACING AND LAG EFFECTS

Five explanations of the spacing effect are particularly relevant to understanding memory formation in an advertising environment: attention, rehearsal, encoding variability, retrieval, and reconstruction. First, we discuss each explanation of the spacing effect. Then, we discuss a meta-analytic strategy for differentiating between the explanations. Serendipitously, these different-iating tests will also be able to inform us about potentially effective advertising practices.

Explanations of the Spacing Effect

Attention Hypothesis. The attention hypothesis attributes the poorer recall in the massed condition to people *voluntarily* paying less attention to P2 (presentation 2) when it occurs shortly after P1 (presentation 1) (Hintzman 1974). People recognize that P2 is repetitive and treat it as a

rest period, waiting for something new to process. The attention hypothesis assumes that people use recognition as a cue that the P2 material is not novel and can safely be ignored. Recognition should decline as the interval between presentations increases; hence, spaced presentations should create better memory for the material.

Rehearsal Hypothesis. The rehearsal hypothesis predicts that a massed presentation schedule inhibits recall because it limits rehearsal of P1 in a P1-P2 presentation sequence (Rundus 1971). The hypothesis assumes short-term memory always contains a set of rehearsal items. The probability of an item being in the rehearsal set decreases as time since its exposure increases. When P2 occurs shortly after P1, it limits the rehearsal of P1 because it is more apt to replace P1 in the rehearsal set. When P2 occurs much later than P1, P1 is rehearsed until it is naturally dropped from the rehearsal set. Since memory is aided by more rehearsal of the stimulus, distributed presentations should create better memory for the material.

Encoding Variability. The encoding variability hypothesis predicts that spaced presentations enhance recall because they allow for the formation of more cue-target associations (Glenberg 1979; Melton 1970). Glenberg (1979) posits that cues can be general (e.g., associations to the learning environment), contextual (e.g., associations to contingent items), and descriptive (e.g., associations to the stimulus). Increasing the amount of time between P1 and P2 creates a greater opportunity for general, contextual, and descriptive cues to change. To the extent that the processing of a stimulus reinforces associations to available cues, a spaced presentation schedule should result in more cue-target associations. Retrieval cues have also been classified as semantic and non-semantic (Challis 1993). Semantic cues can be strong or weak associates of the target stimulus, whereas non-semantic cues are unrelated organizational (e.g., position in list) and

contextual (e.g., learning episode, learning location, font, color, etc.) events.

Retrieval Hypothesis. The retrieval hypothesis predicts that the spacing effect is a positive function of the difficulty of successfully retrieving P1 at P2. When exposed to an event, a person is automatically reminded of other events (Braun and Rubin 1998; Green 1989). In the context of spacing, this means that P2 will serve as a cue for the involuntary retrieval of P1. If P1 is retrieved from long-term memory instead of from working memory at P2, then the person has had the opportunity to engage in retrieval practice. Retrieval practice should enhance the probability the item will be retrieved at a subsequent time.

Semantic Reconstruction (Accessibility) Hypothesis. The semantic reconstruction hypothesis predicts that the spacing effect strengthens to the degree that the stimulus is reconstructed at P2 (Jacoby 1978). The hypothesis assumes that the act of perception requires people to construct a representation of the event and it is easier to retrieve a previous representation of a stimulus than to construct a new representation. If an item is repeated (P2) while the previous representation (P1) is still accessible in short-term memory, then there is no need to construct the event. If the repeated experience is delayed and the P1 representation begins to fade, then it is necessary to go through a reconstruction process at P2 (i.e., construct a representation of the stimulus). The amount of reconstruction is directly related to the amount of P1 that has decayed. As reconstruction is an elaborative process, more reconstruction should lead to better recall (Thios and D'Agostino 1976).

Spacing and Advertising

Differentiating the competing explanations of the spacing effect is important because individual explanations have the potential to make predictions that violate common advertising

wisdom about learning via repetition and to provide insight about how to enhance the effectiveness of individual advertisements and campaigns. For example, one of the oldest findings in memory research is that the relearning of material is easier when the material is presented in the same context (Ebbinghaus 1985). In contrast, both the attention and the encoding variability hypotheses predict it is better to vary the learning context at different exposures to maximize learning. A second example is the rule of thumb that simple ad material is best remembered. In contrast, the retrieval and reconstruction hypotheses predict that memory for complex stimuli will benefit more from repetition and will eventually surpass memory for simple information. Even one of our most fundament learning principles, that involved learning leads to better memory for material, is challenged by retrieval and reconstruction hypotheses predictions that the best learning scenarios are sequences of more and less involved processing. Thus, differentiating among these competing explanations of memory formation has the potential to generate a number of novel hypotheses about learning ad material via repetition.

Our effort to differentiate between the competing explanations of the spacing effect involved a series of ten tests conducted using meta-analysis. An understanding of these tests depends on an understanding of the coding of the content or context variable and the predictions of each individual theory. In the next section, we discuss the data and the meta-analysis. Then we discuss the competing predictions and the results on a hypothesis-by-hypothesis basis.

ANALYSIS

Procedure

An extensive search of the PsychINFO database was performed for the years 1887-2000.

The database was searched using the key terms: spacing effect, lag effect, distributed presentation, spaced presentation. All referenced papers in these published articles were included in the study and the more prolific authors of these published articles were contacted requesting working papers that had not been published. The only inclusion criteria were that the studies investigated a spacing effect and that the statistical information reported in the studies was sufficient to calculate an effect size. Eighty-one articles were collected of which 61 articles reported sufficient statistical information to calculate an effect size. These 61 articles reported 484 tests of the spacing effect when all lags tests from the same study were included. When lag redundancy was removed, 269 tests of the spacing effect remained.

The effect size estimate (Rosenthal 1984) was the product-moment correlation (r) which is the square root of the variance explained by a given variable or combination of variables. Explained variance (EV) can be calculated from any X^2, F, or t statistic with one degree of freedom (EV = X^2 /N; EV = $F/[F + $ df within]; EV = $t^2/[t^2 + $ df]). The correlation coefficient was chosen as the measure of effect size because it is easy to compute from a t or F statistic, it is insensitive to cell sample size differences associated with between-subject tests, and it can be easily interpreted. It is well known that r-values are not normally distributed, so the r-values were transformed into a Fisher's Z_r (e.g., Z_r = .5 [ln(1+r) - ln(1-r)]) prior to performing all statistical tests. Means calculated using the Fisher's Z_r were transformed back into r (e.g., $r = [e^{2Z} -1]/[e^{2Z} + 1]$) for reporting purposes.

A combined z and fail-safe N were calculated for each combined effect size (Rosenthal 1984). A combined z is calculated by expressing t-values as one-tail p values, expressing these p values as z-scores from a standard normal distribution, adding these z-scores, then dividing by the

square root of the number of scores. We report a combined z that is not weighted by respective

sample sizes because it is a more conservative statistic than the weighted combined z. A fail-safe

or "file drawer" N (e.g., $N = [((\sum z_j)^2 / 2.706) - k])$ estimates the number of unpublished studies

with an effect size of zero that would have to exist in order to render the effect insignificant at the

alpha = .05 level. As expected, the meta-analysis of the spacing effect (n = 269) was statistically

significant ($r = .339$; combined $z = 36.83$; fail-safe $N = 148,979$).

The test of the influence of the stimulus content and context factors used 484

observations, but the repeated observations associated with lags were nested within an individual

test of the spacing effect so that, in effect, the tests depended on 269 observations. To perform

these tests, we used a Generalized Estimating Equations (GEE) procedure. The GEE procedure is

recommended when observations are correlated because they are repeated or because they have a

common characteristic (Liang and Zeger 1986). In our data, we had a lag variable that varied

across experiments. The net effect of using the GEE procedure is that spacing experiments with

more levels of lag in the design did not receive additional weight in tests of the manipulated

variables. Note that there was no significant difference in effect size for the six different dependent

measures used in the spacing studies: the percentage correct in a recall task ($n = 216$), the

percentage correct in a frequency estimate ($n = 13$), the percentage correct in a recognition task

($n = 18$), the amount recalled ($n = 8$), the estimated frequency ($n = 13$), and the response latency

($n = 1$); $F(5, 263) = 1.94$, $p > .05$). Thus, we collapsed the data over the type of dependent

measure except for hypothesis that differentiated between free and cued recall tests.

Hypotheses Tests

Ten hypothesis tests were performed. The hypotheses and test results are summarized in

table 1 and the meta-analysis statistics are reported in table 2. A list of the studies used to perform

the meta-analysis is available from the first author.

<center>

****

Insert Table 1 here

****

</center>

The Lag Effect. Investigations into the spacing effect often test multiple spacing intervals

in the same study. The rehearsal, attention, encoding variability, and reconstruction hypotheses

predict a positive linear or logarithmic effect of the amount of the lag (i.e., spacing interval). The

retrieval hypothesis predicts an inverted-U relationship of increasing lag (Bjork 1988). Initially, as

the lag increases, retrieval of P1 becomes more difficult and subsequent memory for the stimulus

increases. At some point, the lag becomes long enough to prohibit retrieval of P1 at P2 and

subsequent memory for the stimulus declines.

Twenty-nine articles provided 83 direct tests of multiple spacing intervals ranging from

five seconds to one week. The 83 tests investigated from two to five spacing intervals. Fifty-one

tests showed a significant, positive influence of lag, 19 showed no influence of lag, five showed a

significant, negative influence of lag, seven showed an inverted-U relationship between effect size

and lag, and one showed an U-shaped relationship between effect size and lag. In an aggregate

test, we set the intercept of all 83 tests to have the same value at the shortest lag in order to

directly assess the effect of going from the shortest to the longest spacing interval. The effect size

averaged .298 at a lag of one, .440 at a lag of two, .543 at a lag of three, .531 at a lag of four and

.565 at a lag of five. The tests for a linear relationship ($F(1, 478) = 49.23$, $p < .01$) and log

relationship ($F(1, 478) = 25.69$, $p < .01$) were both statistically significant, but the test for the

inverted-U quadratic relationship was not significant ($F(1, 477) = 0.89$, $p > .10$).

Verbal / Pictorial Stimuli. Explanations of the spacing effect make different predictions for verbal and pictorial stimuli. The rehearsal hypothesis predicts that pictures are harder to rehearse than words; hence pictures should not benefit from increased rehearsal in the distributed condition and should show a weaker spacing effect. The reconstruction hypothesis predicts an interaction effect based on the type of dependent measure. The reconstruction of pictorial stimuli relies on stimulus features, thus the spacing advantage for pictorial information should be strongest when there is a recall cue that incorporates stimulus features. The reconstruction of verbal information relies on conceptual features, thus the spacing advantage for verbal information should be strongest when there is no recall cue. The remaining three theories make no direct predictions.

The stimuli used in each spacing study were dummy coded as nonsense words, words, sentences, or pictures (see table 2 for mean memory levels and average effect sizes) with the later three types of stimuli contributing to this analysis. Planned contrasts showed that the effect size for the verbal stimuli ($M_{word + sentence} = .338$) was not different from the effect size for pictorial stimuli ($M = .303$; $X^2(1) = 0.17$, $p > .10$), a finding inconsistent with the rehearsal hypothesis. Predictions of the reconstruction hypothesis were investigated after splitting the studies into cued recall and free recall groups. For cued recall, pictures ($M = .430$) showed a stronger spacing effect than verbal stimuli ($M_{word + sentence} = .235$; $X^2(1) = 4.75$, $p < .05$). For free recall, verbal stimuli ($M_{word + sentence} = .397$) showed a stronger spacing effect than pictures ($M = .174$; $X^2(1) = 6.39$, $p < .05$). These results are consistent with the reconstruction hypothesis.

****
Insert Table 2 here
****

Meaningless / Meaningful Stimuli. The rehearsal hypothesis predicts that meaningful stimuli are easier to rehearse and should exhibit a stronger spacing effect. The encoding variability hypothesis predicts that the spacing effect occurs because of the greater number of cues that become associated with the stimulus at P1 and P2; more cues lead to better recall (Postman and Knecht 1983). To understand this prediction, recall that Glenberg (1979) posits that cues can be the related to the learning environment, contingent items, and the descriptive characteristics of the stimulus. Glenberg argues that descriptive cues are stored in semantic memory and that different cues exist at different levels of analysis. Spaced presentations increase the potential for different levels of analysis and reinforce a greater variety of cue-stimulus associations. Since meaningful stimuli have a greater number of descriptive cues, the spacing effect should be stronger for meaningful stimuli.

The retrieval hypothesis predicts that at P2, people attempt to retrieve their memory of P1, a process that should be easier with meaningful stimuli. Hence, there should be a weaker spacing effect for meaningful stimuli. The reconstruction hypothesis predicts that people attempt to retrieve the representation formed at P1 in order to avoid construction of a representation of the stimulus at P2. Access to P1 should be easier with meaningful stimuli; hence, there should be a weaker spacing effect. The attention hypothesis makes no prediction.

The mean memory levels and average effect sizes for meaningful and meaningless stimuli are shown in table 2. Meaningful stimuli ($M = .335$) exhibited a significantly weaker spacing effect

than meaningless stimuli ($M = .509$; $X^2(1) = 20.04$, $p < .05$). These results are consistent with the retrieval and reconstruction hypotheses, but not with the rehearsal and encoding variability hypotheses.

Novel / Familiar Stimuli. The rehearsal, attention, and encoding variability hypotheses predict familiar stimuli will show a stronger spacing effect than novel stimuli. It should be easier to rehearse familiar stimuli; hence, they should benefit more from spaced presentations. With respect to the attention hypothesis, novel stimuli are known to be better at capturing attention than familiar stimuli (Berlyne 1970). To the extent a novel stimulus can better sustain attention at P2 in the massed condition, the size of the spacing effect will be limited. Encoding variability theory predicts a stronger spacing effect for familiar stimuli because familiar stimuli have more pre-existing associates. Spaced presentations allow the stimulus to be processed at more levels of analysis, creating opportunities for a wider set of cues to become more strongly associated with a stimulus and, hence, improve recall.

In contrast to the above explanations, the retrieval and reconstruction hypotheses predict that the spacing effect should be stronger for novel stimuli. The retrieval hypothesis prediction is based on the assumption that novel stimuli are more difficult to retrieve at P2. The reconstruction hypothesis prediction is based on the assumption that novel stimuli are harder to retrieve and are more likely to be reconstructed at P2.

The mean memory levels and average effect sizes for novel and familiar stimuli are shown in table 2. The effect size for novel stimuli ($M = .424$) was larger than the effect size for familiar stimuli ($M = .347$), but this difference was not statistically significant ($X^2(1) = 2.72$, $p > .10$). We were concerned that this test was non-diagnostic due to the lack of power created by the small

number of observations in the novel stimulus condition (e.g., $n = 12$). An additional test was conducted using the standardized mean performance data reported in table 2. The difference in the average proportion of correct responses was greater for studies using novel stimuli ($M_{\text{massed}} = .411$, $M_{\text{distributed}} = .575$) than familiar stimuli ($M_{\text{massed}} = .404$, $M_{\text{distributed}} = .511$; ($F(1, 214) = 4.78$, $p < .05$). The direction of the effect size differences in the first test and the statistical significance of the second test are consistent with the predictions of the retrieval and reconstruction hypotheses.

Complex / Simple Stimuli. Stimuli can be simple (e.g., the word "cat"), structurally complex (e.g., the sentence "The cat is on the red brick wall.") or semantically complex (e.g., homographs) (Challis 1993; Green 1989). The rehearsal hypothesis predicts that structurally and semantically complex stimuli should be more difficult to rehearse than simple ones. The lower rehearsal rate in the spaced condition should lead to a weaker spacing effect for complex stimuli. In contrast, the encoding variability hypothesis predicts that a complex stimulus has more potential associations and, since spaced presentations allow for a greater variety of these associations to be reinforced, there should be a stronger spacing effect for complex stimuli. The retrieval and reconstruction hypotheses further differentiate between spacing effects for structurally and semantically complex stimuli. It should be more difficult to retrieve or reconstruct a semantically complex P1 stimulus at P2, and this additional processing should lead to a stronger spacing effect than for simple stimuli. In general, retrieval does not depend on structural cues, so variability in structural complexity should not influence the spacing effect.

The mean memory levels and average effect sizes for simple, structurally complex, and semantically complex stimuli are shown in table 2. A test for differences in the average effect size

across the three different types of stimulus complexity was statistically significant (X^2 (2) = 23.13, $p < .05$). The average effect size for structurally complex stimuli ($M = .330$) was not significantly different than the average effect size for simple stimuli ($M = .325$; $X^2(1) = 1.56$, $p > .05$). The average effect size for semantically complex stimuli ($M = .586$) was significantly greater than the average effect size for simple stimuli ($M = .325$; $X^2(1) = 19.95$, $p < .05$). These results are consistent with the retrieval and reconstruction hypotheses, partially consistent with the encoding variability hypotheses, and inconsistent with the rehearsal hypothesis.

      Uni-modal / Bi-modal Presentation Format. Stimuli can be presented using a uni-modal format (e.g., only visual presentation, only auditory presentation) or a bi-modal format (combined visual and auditory presentation). The attention, retrieval, and reconstruction hypothesis all predict a smaller spacing effect of a bi-modal format. The attention hypothesis assumes that a bi-modal presentation should be able to sustain attention to P2 in a massed presentation context; hence the spacing effect should be weaker than for a uni-modal format. The retrieval hypothesis assumes that a bi-modal presentation of distributed stimuli will make the retrieval of P1 easier at P2 because there will be more retrieval cues. The retrieval hypothesis predicts that the spacing effect is enhanced by the difficulty of successful retrieval, so the bi-modal presentation should show a weaker spacing effect. The reconstruction hypothesis assumes that a bi-modal presentation format will make it easier to retrieve P1 and avoid the construction of a representation at P2; hence, the bi-modal presentation should show a weaker spacing effect. In contrast, the encoding variability hypothesis predicts that a bi-modal presentation should provide more structural cues (verbal, visual) during training, create more opportunities for cues to become associated with a stimulus, and result in a stronger spacing effect.

The mean memory levels and average effect sizes for uni-modal visual, uni-modal auditory, and bi-modal stimuli are shown in table 2. The average effect sizes for visual ($M = .290$), auditory ($M = .449$), and multi-modal ($M = .302$) stimuli were significantly different ($X^2 (2) = 30.22, p < .05$). However, with the uni-modal formats combined ($M = .347$), the average effect size did not differ significantly from the average effect size for multi-modal stimuli ($M = .302$; $X^2(1) = 0.51, p > .10$). These results are inconsistent with the predictions of all of the hypotheses.

Unrelated / Related Cues in a Paired-Associate Learning Task. Stimuli can be learned in a paired associate task in which the cue is unrelated (e.g., wood – snow), structurally related (e.g., crow-snow), or semantically related (e.g., ice – snow). The rehearsal hypothesis predicts that related cues should make cue-target associations easier to rehearse. Sustaining rehearsal between P1 and P2 should lead to a stronger spacing effect. The encoding variability hypothesis predicts no influence of cue relatedness. Experiencing the same cue at P1 and P2 limits the degree to which other cues can become associated with the target stimulus. The retrieval and reconstruction hypotheses predict that it will be more difficult to retrieve or reconstruct a stimulus as the target and cue become less related. This additional processing should lead to a stronger spacing effect for less related targets than for more related targets.

The mean memory levels and average effect sizes for unrelated, structurally related, and semantically related cues are shown in table 2. A test for differences in the average effect size across different types of cues was not statistically significant ($X^2(2) = 1.35, p > .10$). Also, in a separate test, the average effect size for unrelated cues ($M = .297$) did not differ from the effect size for related cues ($M = .402$; $X^2(2) = 1.21, p > .10$). However, since the small number of studies investigating cue relatedness meant low statistical power, we ran an additional test using

the standardized mean memory performance data reported in table 2. The improvement in memory for structurally related ($M_{\text{massed}} = .466$, $M_{\text{distributed}} = .665$) and semantically related ($M_{\text{massed}} = .415$, $M_{\text{distributed}} = .539$) cues was significantly greater than the memory improvement for unrelated cues ($M_{\text{massed}} = .455$, $M_{\text{distributed}} = .525$; $F(1, 28) = 7.53$, $p < .05$). This finding is consistent with the retrieval and reconstruction hypotheses, whereas the statistically insignificant test of the spacing effect is consistent with the encoding variability hypothesis. However, we note that the one direct, experimental test of this hypothesis (Johnston et al. 1972) found no difference between a semantically related cue (sports-fan) and an unrelated cue (engagement-fan).

Intentional /Incidental Learning. People can process material with the intent to learn it or be exposed incidentally. All of the explanations predict a larger spacing effect when material is processed intentionally, but for different reasons: varying the spacing of the stimuli will have a larger influence when the respondent is actively rehearsing the material (rehearsal hypothesis); people engaged in intentional learning are more likely ignore P2 in the massed condition because they recognize the presentation as redundant (attention hypothesis); intentional learning should reinforce more associations at P1 and P2 than incidental learning (encoding variability hypothesis); intentional processing will create a stronger spacing effect because it encourages a person to actively attempt to retrieve P1 (retrieval hypothesis); intentional processing will create a stronger spacing effect because P2 will be elaborated in the attempt to retrieve P1, and when not successful, to construct a representation of P2 (reconstruction hypothesis).

The mean memory levels and the average effect sizes for intentional and incidental learning are shown in table 2. The average effect size for intentional processing ($M = .352$) was larger than for incidental processing ($M = .236$; $X^2(1) = 9.24$, $p < .05$). These results confirm the predictions

of all of the hypotheses.

Three experiments have manipulated the learning goal for a direct test of its influence on the spacing effect. Greene (1989) found that intentional learning influenced free recall and cued recall in different ways. For free recall, spacing effects were unaffected by the degree of intentional learning, but a spacing effect was obtained for both the intentional and incidental conditions. However, for cued recall, the spacing effect was eliminated under incidental learning conditions. Challis (1993) found that a spacing schedule improved recall in cued-memory tests under intentional and incidental-semantic learning conditions but not in an incidental learning condition that encouraged graphemic (surface) processing of the stimuli. However, in a test of the retrieval hypothesis, Braun and Rubin (1998) found a significant spacing effect in *both* cued and free recall tests under both intentional and incidental learning conditions.

Isolated /Embedded Stimuli. Stimuli can be presented in isolation (i.e., one word or image at a time) or as part of a larger whole (e.g., a target word in a sentence or a brand name in an ad). The rehearsal, attention, and encoding variability hypotheses predict that isolated stimuli should show a stronger spacing effect than embedded stimuli, whereas the retrieval and reconstruction hypotheses predicted the opposite. The rehearsal hypothesis predicts that embedded stimuli will compete with material from the embedding context for access to short term rehearsal and, hence, the spacing effect will be weaker than when the stimuli are isolated. The attention hypothesis posits that isolated stimuli should be more noticeable at P1 (the Von Restorff effect), receive less attention at P2, and thus show a stronger spacing effect compared to the embedded stimuli. The encoding variability hypothesis predicts that a stimulus surrounded by other items should have an opportunity for variable encoding, but to the extent this surrounding context is the same at P1 and

P2, this variable encoding should remain constant. In other words, there is no advantage to spacing the presentations. Thus, there should be greater opportunity for variable encoding when the stimulus is presented in isolation.

The retrieval and reconstruction hypotheses predict isolated stimuli should show a weaker spacing effect than embedded stimuli. The retrieval hypothesis prediction is based on the assumption that it is easier to retrieve an isolated P1 stimulus at P2. The reconstruction hypothesis prediction is based on the assumption that it is easier to retrieve an isolated P1 stimulus and avoid reconstruction of the stimulus at P2; hence, the isolated stimuli should show a weaker spacing effect than the embedded stimuli.

The mean memory levels and the average effect sizes for stimuli presented in an isolated or embedded context are shown in table 2. The difference between the average effect size for the isolated ($M = .333$) and embedded ($M = .371$) learning contexts was not significant ($X^2(1) = 0.26$, $p > .10$). These results are inconsistent with all of the hypotheses.

Simple / Complex Intervening Material. The material that is processed between repeated presentations can be simple (e.g., nothing, mono-syllabic words), structurally complex (e.g., multi-syllabic words, sentences), or semantically complex (e.g., text passages, sentences in a story line). The rehearsal hypothesis prediction is opposite to those of the encoding variability, retrieval, and reconstruction explanations. It assumes that semantically and structurally complex intervening material is more difficult to rehearse and, thus, interferes with the rehearsal of the target material. As a consequence, the target material becomes less sensitive to the spacing manipulation. In contrast, the encoding variability hypothesis assumes that the surrounding material is a primary contextual cue and complex surrounding material will engender a stronger spacing effect. The

retrieval and reconstruction hypotheses make more precise predictions than the encoding variability explanation. The former predicts that semantically complex intervening material should make retrieval of P1 more difficult at P2, thus resulting in a stronger spacing effect relative to the simple and structurally complex intervening material conditions. The reconstruction hypothesis also predicts a stronger spacing effect for semantically complex intervening material since this material should make it more difficult to retrieve P1 and force reconstruction of the stimulus representation at P2.

The mean memory levels and average effect sizes for simple and complex intervening material are shown in table 2. A test for differences in the average effect size across different types of intervening stimulus complexity was statistically significant ($X^2(2) = 6.78$, $p < .05$). The effect size for semantically complex intervening stimuli ($M = .419$) was significantly greater than when simple intervening stimuli ($M = .331$) or structurally complex intervening stimuli ($M = .327$) were used ($X^2(1) = 5.55$, $p < .05$). The effect size for structurally complex intervening stimuli ($M = .327$) was not significantly greater than for simple intervening stimuli ($M = .331$) were used ($X^2(1) = 2.30$, $p > .10$). These results are consistent with the predictions of the retrieval and reconstruction hypotheses.

Limitations of the Meta-Analysis

Our results and any implied conclusions share limitations common to all meta-analyses. First, although the data come from experimental studies, our results are essentially based on a "meta correlation" of the size of the spacing effect and the variables coded across these experimental studies. Except for the four noted experiments, the studies used in the meta-analysis included no direct, non-confounded tests of the hypotheses. Second, the results are based only on

the data available from a "natural" (i.e., non-designed) experiment (Farley, Lehmann, and Sawyer 1995). That is, although one can conceive a matrix of studies that includes all combinations of the nine factors, the 248 studies in our analysis obviously fall far short of the 1,152 cells that would be required to construct this matrix. The limited number of studies within this matrix implies the possible presence of confounding variables. Third, some of our tests, such as those investigating semantic complexity, are underpowered due to too few studies including the variable in question. Fourth, this meta-analysis has not weighted individual studies on the basis of their quality. Research like ours has been labeled a "bare-bones" meta-analysis, since it considers only the quantitative factors of (1) levels of the potential moderating factor, (2) sample size, and (3) sample effect sizes. Hunter and Schmidt (1990) identified thirteen qualities, such as measurement error, that can distinguish the potential imperfections of a study. However, we preferred to confine our analyses to substantive and procedural variables that could be coded with considerably less controversy, rather than to introduce our biases about study quality.

DISCUSSION

There were two goals for this research study. The first goal was to gain insight into the processes responsible for the memory improvement that results from repetition. We used a meta-analysis to compare five theories: the rehearsal, attention, encoding variability, retrieval, and reconstruction hypotheses. The results are most consistent with the retrieval and reconstruction hypotheses. The retrieval (six successful predictions) and reconstruction (eight successful predictions) theories perform best, followed by encoding variability with three correct predictions. This result is surprising given that voluntary attention (e.g., Malaviya and Sternthal 1997) and

encoding variability (e.g., Schumann et al. 1990; Singh et al. 1994; Unnava and Burnkrant 1991) have been the dominant explanations of spacing effects within the consumer behavior literature. Thus, even though the retrieval and reconstruction theories are well established in the verbal learning literature, they have not been used to understand repetition and advertising memory effects.

The retrieval and reconstruction hypotheses have interesting implications about how to use repetition to improve memory of advertising material. The retrieval explanation emphasizes the importance of obtaining a strong initial encoding of a stimulus at P1 but sufficiently limiting retrieval cues at P2 such that the retrieval is difficult. The reconstruction explanation stresses the importance of a strong, partial encoding at P1 in order to encourage elaboration at P2 as a representation of the stimulus is constructed. Thus, both explanations argue that an optimal repetition strategy should encourage incidental processing during one presentation and intentional processing in the other but differ about the optimal order of these two types of processing. Conceptually, this means that a combination of elaborated and incidental processing of the same advertising material may result in better memory for the material then two occasions of elaborated processing. Practically, this means that a more effective repetition strategy may include a combination of involving media (e.g., television commercials) and less involving media (e.g., billboards, product placements) or messages that vary in the level of involvement – for example, complex versus simple; long versus short; hard sell versus soft sell (Ray and Sawyer 1971a). This variability in involvement does not have to occur across ads. Varying the depth and breadth of encoding at P1 and P2 within a specific ad may make the ad more effective at promoting memory for the material. It is interesting to note that Krugman's (1970) discussion of the processing of

three advertising exposures, with a first exposure evoking a "what is it?" response, the second exposure causing a "what of it?" reaction, and a third exposure reinforcing previous responses, implies that consumers learn when there is incidental processing at P1 and elaborated processing at P2. Thus, Krugman's model is consistent with the learning mechanism described by the reconstruction explanation.

Although the advertising context can be used to vary the depth of processing at P1 and P2, our results suggest that stimulus characteristics can be manipulated to accomplish the same goals. For example, the retrieval and reconstruction explanations were the only ones to propose the largest effect found in the meta-analysis --- the greater "depth of processing" afforded by increased semantic complexity enhances the spacing effect. These two explanations were also the only ones to predict that more semantically complex intervening material would result in a stronger spacing effect. Thus, if empirical support for differentiating predictions is an important criterion for identifying promising theories, then the retrieval and reconstruction hypotheses fare well in the meta-analysis.

Moderators of the Spacing Effect

The ten hypotheses used to discriminate between the explanations of the spacing effect investigated a number of content and context variables that are important in advertising. Some of the content and context variables have garnered a small amount of consumer behavior research but have lacked the theoretical underpinnings needed to make a conceptual advance. In this section, we attempt to integrate the findings of the meta-analysis, the existing consumer behavior literature, and the assumptions of the retrieval and reconstruction hypotheses to create some interesting insights and propositions about repetition and memory. At a minimum, these

observations might generate a promising research agenda and might be considered working rules of thumb for advertising practitioners.

Repetition and Memory. It is well known that repetition of advertising is an effective learning tool. Laboratory (e.g., Singh et al. 1994) and field (Zielske 1959) evidence show that memory for repeated material improves as the time between presentations of advertising material increases, especially when there is a delay between the second presentation of the stimulus and the memory test (Singh et al. 1994). Yet, there is also evidence that repeated exposures are ineffective, a somewhat counterintuitive finding. Appel (1971) and Blair (2000) present field evidence that memory for a second (or nth) exposure to an ad declines as the time between the first and second exposure increases, even if the recall measurement is administered soon after the second exposure. They also find that the decline in recall is not as severe if the initial exposure to the ad resulted in a more memorable or persuasive experience. These results are consistent with the retrieval hypothesis. To the extent the initial exposure is engaging, a strong memory trace is created. As the time between P1 and P2 increases, the number of people that can retrieve this trace at P2 declines; hence, the benefit of the first exposure declines. Yet, the P1 memory failure will be slower for initially engaging ads. Note that if engaging ads simply capture attention at P2, memory at test should be a function of the time since P2, not the time between P1 and P2, because everyone is assumed to process the ad at P2. Our proposed explanation of a well-documented, counterintuitive field finding would benefit from laboratory investigation.

Pictorial / Verbal Stimuli. The recall of pictures has been found to be longer lasting than the recall of words (Gardner and Houston 1986). Our analysis indicates that spacing should benefit pictorial stimuli more than verbal stimuli when cued recall tests are used. Consistent with

this finding, Keller (1987) found that a picture in a point of purchase display from a memorable ad campaign facilitated the cued recall of ad claims. On the other hand, our analysis shows that in situations where consumers might engage in memory-based choice, verbal ads would be more sensitive to repetition–based advertising strategies.

Ray, Sawyer, and Strong (1971) found that the repetition of color ads resulted in greater increases in ad recall than the repetition of otherwise equivalent black and white ads. However, when the measure of memory was the depth of recall, repetition of black and white ads was more effective than the repetition of color ads. These data suggest that repeated exposure might have facilitated recall for different types of information given different retrieval / reconstruction cues (e.g., color might have been an effective retrieval cue for ad identity, whereas brand name and product category information might have been effective retrieval cues for ad content). These results are consistent with the retrieval and reconstruction explanations of the spacing effect.

Stimulus Meaning. Keller, Heckler, and Houston (1998) suggest that it is better to use meaningful brand names when selecting a brand name for a new product. Our results show that meaningful stimuli are better recalled than meaningless stimuli, supporting their view. However, our data also suggest that memory for initially meaningless brand names will benefit more from distributed exposures. In other words, appropriately designed ad schedules will more effectively build memory for an initially meaningless brand name. This may be especially useful to advertisers who need to use brand names that are free of preexisting associations.

If advertisers do use initially meaningless brand names or product logos, it is interesting to hypothesize about how repetition can encourage the development of meaningful brand associations (e.g., benefits, usage situations). Conventional wisdom assumes that repetition is

better than no repetition and elaborated processing is better than incidental processing. Yet, the retrieval and reconstruction explanations argue that associated meanings should be primary in some exposures, but secondary in other exposures and that presentations of these two types of ads should be spaced. The hypothesis that primary information will encourage reconstruction of secondary information in previous ads or that secondary information will encourage retrieval of primary information from previous ads is novel. The prediction that either of these ad presentation strategies will result in stronger long-term memory than a strategy that relies on elaborated processing of an ad at each presentation is intriguing and worthy of empirical investigation.

Stimulus Complexity. Some research has examined the use of complex ads and brand names on the recall of ad information. McQuarrie and Mick (1992) varied the degree of ad resonance, a type of complex homograph, and found that ads with resonance produced better recall of ad headlines as well as greater liking for both the ad and the brand. Park, Jun, and Shocker (1996) examined the effectiveness of complex, composite brand names on attitudes toward a brand extension context. Stevenson, Bruner, and Kumar (2000) found that simple web page backgrounds were more effective than complex backgrounds in positively affecting attitude. However, in a verbal learning study that examined how complexity impacted memory, Gilbert (1998) found that using a complex, composite brand name facilitated greater processing and recall. Gilbert reasoned that complex, composite brand names were more likely to attract attention, encourage involvement, and facilitate elaboration at the time of encoding. The retrieval explanation would argue that this recall could be further enhanced if subsequent exposures to the brand name were distributed in time and less involving. Making subsequent exposures to the complex, composite brand name involving would limit improvements in memory for the brand, a

counterintuitive prediction.

Varied Ad Executions. Several researchers have argued that varied ad executions are necessary in order to enhance recall. Berlyne's (1970) two-factor theory hypothesized that the effectiveness of repeated exposures depend on whether the negative effects of tedium outweigh the positive effects of reduced uncertainty. He hypothesized that more complex and varied exposures would stave off tedium, encourage attention to subsequent exposures, and facilitate memory. Berlyne's concerns are echoed in media planners' concerns about ad *wear-out*; a loss in ad effectiveness at a high number of exposures owing to boredom, inattention, or irritation (Naik, Mantrala, and Sawyer 1998; Pechmann and Stewart 1989). Consistent with a spacing effect, advertisers presume that the effectiveness of ads decline with high amounts of concentrated repetition and that an ad may alleviate wear-out by using a more spaced exposure schedule. In some cases, the effectiveness of an ad may be improved if it is retired and then reintroduced after a considerable length of time (Naik et al. 1998).

Consistent with Berlyne and the notion of wear-out in media planning, repetition of similar but non-identical ads has been shown to result in higher recall than repetition of identical ads (e.g., Unnava and Burnkrant 1991). Also, Singh et al. (1994) concluded that advertisers should vary the context in which ads appear in order to maximize recall. These two studies used encoding variability theory to explain the underlying processes and suggest the resulting practical implications. Varying ad content, length, or the context in which ad appears should lead to an increased number of retrieval routes. However, our analysis indicates that enhanced recall will more likely result from "deeper levels of processing" which should lead to one strong retrieval route instead of the many retrieval routes predicted by encoding variability. Moreover, the

benefits of this variation strategy would be greater for massed exposures than for spaced repetitions. One major implication for the marketer is that strategies leading to a "deeper processing" of ad content during one of the two exposures should be a more effective memory retention strategy.

Summary

A great amount of research has investigated the finding that distributed exposures of a stimulus lead to better memory than massed exposures of a stimulus. Ninety-seven studies and 269 independent tests of the spacing effect were used to differentiate between five theoretical explanations. Each explanation of the spacing effect was used to generate hypotheses about ten potential moderating variables. Two explanations, the retrieval hypothesis and reconstruction hypothesis, are most consistent with the results. We propose a program for future research that should provide a better understanding of the spacing effect and help identify scheduling strategies that can maximize the benefits of repeated advertising exposure.

REFERENCES

Appel, Valentine (1971), "On Advertising Wear Out," *Journal of Advertising Research*, 11 (February), 11-13.

Berlyne, Donald E. (1970), "Novelty, Complexity, and Hedonic Value,≅ *Perception and Psychophysics*, 8 (November), 279-286.

Bjork, Robert A. (1988), "Retrieval Practice and the Maintenance of Knowledge," in *Practical Aspects of Memory II*, ed. Michael M. Gruneberg, Peter E. Morris, and Robert N. Sykes, London: Wiley, 396-401.

Blair, Margaret Henderson (2000), "An Empirical Investigation of Advertising Wearin and Wearout," *Journal of Advertising Research*, 40 (November), 95-100.

Braun, Kelly Ann and David C. Rubin (1998), "A Retrieval Model of the Spacing Effect," *Memory*, 6 (January), 37-65.

Challis, Bradford H. (1993), "Spacing Effects on Cued-Memory Tests Depend on Level of Processing," *Journal of Experimental Psychology: Learning, Memory and Cognition*, 19 (March), 389-396.

Cornell, Edward H. (1980), "Distributed Study Facilitates Infants' Delayed Recognition Memory," *Memory and Cognition*, 8 (November), 539-542.

Craig, C. Samuel, Brian Sternthal, and Clark Leavitt (1976), "Advertising Wearout: An Experimental Analysis," *Journal of Marketing Research*, 13 (November), 365-372.

_____ (1996), "Distributing and Managing the Conditions of Encoding and Practice," in *Handbook of Perception and Cognition: Memory*, ed. Elizabeth Lignon Bjork and Robert A. Bjork, California: Academic Press, 317-344.

Ebbinghaus, H. (1913), *Memory* (H. A. Ruger and C. E. Bussenius, Trans.). New York: Teachers College. (Original work published 1885) (Paperback ed., New York: Dover, 164).

Farley, John U., Donald R. Lehmann, and Alan G. Sawyer (1995), ""Empirical Marketing Generalization Using Meta-Analysis," *Marketing Science*, 14 (September), G36-G46.

Gardner, Meryl P. and Michael J. Houston (1986), "The Effects of Verbal and Visual Components of Retail Communications," *Journal of Retailing*, 62 (Spring), 64-78.

Gilbert, Robert J. (1998), "The Relative Communications Effectiveness of Composite Brand Name Strategies," *Dissertation Abstracts International*, 59 (4-A), 1257: University of Pittsburgh.

Glenberg, Arthur M. (1979), "Component-levels Theory of the Effects of Spacing of Repetitions on Recall and Recognition," *Memory and Cognition*, 7 (March), 95-112.

_____ and Thomas S. Lehmann (1980), "Spacing Repetitions Over 1 Week," *Memory and Cognition*, 8 (November), 528-538.

Glover, John A. and Alice J. Corkill (1987), "Influence of Paraphrased Repetitions on the Spacing Effect," *Journal of Educational Psychology*, 79 (June), 198-199.

Greenberg, Allan and Charles Suttoni (1973), "Television Commercial Wearout," *Journal of Advertising Research*, 13 (October), 47-54.

Greene, Robert L. (1989), "Spacing Effects in Memory: Evidence for a Two Process Account," *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15 (May), 371-377.

Hintzman, Douglas L. (1974), "Theoretical Implications of the Spacing Effect," in *Theories in Cognitive Psychology: The Loyola Symposium*, ed. Robert L. Solso, Hillsdale, NJ: Erlbaum, 77-99.

_____ and Miriam K. Rogers (1973), "Spacing Effects in Picture Memory," *Memory and*

Cognition, 1 (October), 430-434.

Hunter, John E. and Frank L. Schmidt (1990), *Methods of Meta-analysis: Correcting Error and Bias in Research Findings*, Newbury Park: Sage.

Jacoby, Larry L. (1974), "The Role of Mental Contiguity in Memory: Registration and Retrieval Effects," *Journal of Verbal Learning and Verbal Behavior*, 13 (October), 483-496.

_____ (1978), "On Interpreting The Effects of Repetition: Solving a Problem Versus Remembering a Solution," *Journal of Verbal Learning and Verbal Behavior*, 17 (December), 649-667.

Keller, Kevin L. (1987), "Memory Factors in Advertising: The Effect of Advertising Retrieval Cues on Brand Evaluations," *Journal of Consumer Research*, 14 (December), 316-333.

_____, Susan E. Heckler and Michael J. Houston (1998), "The Effects of Brand Name Suggestiveness on Advertising Recall," *Journal of Marketing*, 62 (January), 48-59.

Krugman, Herbert (1970), "Why Three Exposures May Be Enough," *Journal of Advertising Research*, 12 (November-December), 11-14.

Liang, K.-Y. and Zeger, S. L. (1986), "Longitudinal Data Analysis Using Generalized Linear Models," *Biometrika*, 73, 13-22.

Malaviya, Prashant and Brian Sternthal (1997), "The Persuasive Impact of Message Spacing," *Journal of Consumer Psychology*, 6 (3), 233-255.

McQuarrie, Edward F. and David G. Mick (1992), "On Resonance: A Critical Pluralistic Inquiry into Advertising Rhetoric," *Journal of Consumer Research*, 19 (September), 180-197.

Melton, Arthur W. (1970), "The Situation with Respect of the Spacing of Repetitions and Memory," *Journal of Verbal Learning and Verbal Behavior*, 9 (October), 596-606.

Naik, Prasad A., Murali K. Mantrala, and Alan G. Sawyer (1998), "Planning Media Schedules in the Presence of Dynamic Advertising Quality," *Marketing Science*, 17 (3), 214-235.

Park, C. Whan, Sung Youl Jun, and Allan D. Shocker (1996), "Composite Branding Alliances: An Investigation of Extension and Feedback Effects," *Journal of Marketing Research*, 33 (November), 453-466.

Pechmann, Cornelia and David W. Stewart (1989), "Advertising Repetition: A Critical review of Wearin and Wearout," in *Current Issues and Research in Advertising*, Vol. 11 (2), ed. James H. Leigh and Claude R. Martin, Ann Arbor, MI: University of Michigan, 285-330.

Postman, Leo and Kitty Knecht (1983), "Encoding Variability and Retention," *Journal of Verbal Learning & Verbal Behavior*, 22 (April), 133-152.

Ray, Michael L. and Alan G. Sawyer (1971a), "Repetition in Media Models: A Laboratory Technique," *Journal of Marketing Research*, 8 (February), 20-29.

_____ and _____ (1971b), "Behavioral Measurement for Marketing Models: Estimating the Repetition Function for Advertising Media Models," *Management Science*, 19 (December), 73-89.

_____, _____, and Edward C. Strong (1971), "Frequency Effects Revisited," *Journal of Advertising Research*, 11 (January-February), 14-20.

Rethans, Arno, John L. Swasy, and Lawrence J. Marks (1986), "Effects of Television, Commercial Repetition, Receiver Knowledge, and Commercial Length: A Test of the Two-Factor Model," Journal of Marketing Research, 23 (February), 50-61.

Rosenthal, Robert (1994), "Parametric Measures of Effect Size," in *The Handbook of Research Synthesis*, ed. Harris Cooper and Larry V. Hedges, New York: Sage, 231-244.

Rothkopf, Ernst Z. and Esther U. Coke (1966), "Variations in Phrasing, Repetition Intervals, and the Recall of Sentence Material," *Journal of Verbal Learning and Verbal Behavior*, 5, 86-91.

Rundus, Dewey (1971), "Analysis of Rehearsal Processes in Free Recall," *Journal of Experimental Psychology*, 89 (July), 63-77.

Sawyer, Alan G. (1981), ""Repetition, Cognitive Response, and Persuasion," in *Cognitive Responses to Persuasion*, ed. Richard Petty, Tom Ostrum, and Timothy Brock, _____, 237-262.

Shaughnessy, John J. (1976), "Persistence of the Spacing Effect in Free Recall Under Varying Incidental Learning Conditions," *Memory and Cognition*, 4 (July), 369-377.

Schumann, David W., Richard E. Petty, and D. Scott Clemons (1990), "Predicting the Effectiveness of Different Strategies of Advertising Variation: A Test of the Repetition Variation Hypothesis," *Journal of Consumer Research*, 17 (September), 192-202.

Singh, Surendra N., Sanjay Mishra, Neeli Bendapudi and Denise Linville (1994), "Enhancing Memory of Television Commercials Through Message Spacing," *Journal of Marketing Research*, 31 (August), 384-392.

Stevenson, Julie S., Gordon C. Bruner II, and Anand Kumar (2000), "Webpage Background and Viewer Attitudes," *Journal of Advertising Research*, 40 (January-April), 29-34.

Thios, Samuel J. and Paul R. D'Agostino (1976), "Effects of Repetition as a Function of Study-Phase Retrieval," *Journal of Verbal Learning and Verbal Behavior*, 15 (October), 529-536.

Unnava, H. Rao and Robert E. Burnkrant (1991), "Effects of Repeating Varied Ad Executions on Brand Name Memory," *Journal of Marketing Research*, 28 (November), 406-416.

Zielske, Hubert A. (1959), "The Remembering and Forgetting of Advertising," *Journal of Marketing*, 23, 239-243.

TABLE 1

TESTING COMPETING ACCOUNTS OF THE SPACING EFFECT

Hypotheses	Rehearsal	Attention	Encoding Variability	Retrieval	Recon/Access	Level	r	X^2
1. There is a ____ relationship between lag and the magnitude of the spacing effect.	**Log**[1]	**Log**	**Log**	Inverted-U	**Log**	Log[2]		25.67[a]
						Inverted-U[2]		0.89
2. The spacing effect for verbal stimuli will be ____ than for pictorial stimuli.	Stronger				**Cued: Weaker**	Cued: Verbal	.235	
						Cued: Pictorial	.430	4.75[a]
					Free: Stronger	Free: Verbal	.397	
						Free: Pictorial	.174	6.39[a]
3. The spacing effect for meaningful stimuli will be ____ than for meaningless stimuli.	Stronger		Stronger	**Weaker**	**Weaker**	Meaningful	.335	
						Meaningless	.509	20.04[a]
4. The spacing effect for familiar stimuli will be ____ than for novel stimuli.	Stronger	Stronger	Stronger	**Weaker**[3]	**Weaker**	Familiar	.338	
						Novel	.424	2.72
5. The spacing effect for structurally/ semantically complex stimuli will be ____ than for simple stimuli.	Weaker		Stronger	**Stronger (Semantically Only)**	**Stronger (Semantically Only)**	Struc. Complex	.330	
						Simple	.325	1.56
						Seman. Complex	.586	
						Simple	.325	19.95[a]
6. The spacing effect for bi-modal stimuli will be ____ than for uni-modal stimuli.		Stronger	Weaker	Stronger	Stronger	Bi-modal	.302	
						Uni-modal	.347	0.51
7. The spacing effect for related cues will be ____ than for unrelated cues.	Stronger		**Null**	**Weaker**[3]	**Weaker**	Related	.404	
						Unrelated	.297	1.21
8. Intentional processing will result in a ____ spacing effect than incidental processing.	**Stronger**	**Stronger**	**Stronger**	**Stronger**	**Stronger**	Intentional	.352	
						Incidental	.236	9.24[a]
9. The spacing effect for isolated stimuli will be ____ than for embedded stimuli.	Stronger	Stronger	Stronger	Weaker	Weaker	Insolated	.333	
						Embedded	.371	.26
10. The spacing effect for structurally/ semantically complex intervening material will be ____ than for simple intervening material.	Weaker		Stronger	**Stronger (Semantically Only)**	**Stronger (Semantically Only)**	Struc. Complex	.327	
						Simple	.331	2.30
						Seman. Complex	.419	
						Simple	.331	5.55[a]

Note. 1 – Predictions listed in **bold** are supported by the data. 2 – F statistic reported. 3 – Additional test data reported in manuscript.
a $p < .05$.

TABLE 2

INFLUENCE OF STIMULUS FACTORS ON SIZE OF SPACING EFFECT

Spacing	Number of Qualifying Cases	Combined Effect Size (r)	Combined Z	p	Fail-safe N	% Recall Massed	% Recall Spaced
Stimulus Form							
Nonsense Word	14	0.454	8.35	< .001	347	0.333	0.490
Word	189	0.330	30.06	< .001	62909	0.405	0.513
Sentence	40	0.378	17.49	< .001	4480	0.336	0.459
Picture	24	0.303	10.91	< .001	1031	0.551	0.651
Stimulus Meaningfulness							
Meaningless	9	0.335	6.93	< .001	151	0.267	0.444
Meaningful	256	0.509	36.44	< .001	125378	0.411	0.518
Stimulus Familiarity							
Familiar	253	0.338	37.25	< .001	129469	0.404	0.511
Novel	12	0.424	5.83	< .001	139	0.411	0.575
Stimulus Complexity							
Simple	200	0.325	30.04	< .001	66518	0.406	0.513
Structurally Complex	57	0.330	18.38	< .001	7060	0.413	0.512
Semantically Complex	12	0.586	14.08	< .001	867	0.341	0.526
Stimulus Variety							
Uni-modal: Visual	134	0.290	21.76	< .001	23311	0.395	0.485
Uni-modal: Auditory	75	0.449	26.00	< .001	18656	0.391	0.550
Bi-modal	60	0.302	16.42	< .001	5920	0.433	0.532
Cue Relatedness							
Unrelated	19	0.297	6.65	< .001	292	0.455	0.525
Structurally Related	4	0.455	9.02	< .001	116	0.460	0.665
Semantically Related	15	0.391	8.87	< .001	421	0.415	0.539
Learning Goal							
Intentional	238	0.352	29.83	< .001	78050	0.406	0.524
Incidental	31	0.236	10.06	< .001	1128	0.383	0.440
Presentation Context							
Isolated	228	0.333	33.68	< .001	95365	0.410	0.521
Embedded	41	0.371	16.59	< .001	4129	0.338	0.447
Intervening Material							
Simple	215	0.331	32.39	< .001	83149	0.397	0.510
Structurally Complex	39	0.327	16.82	< .001	4040	0.502	0.587
Semantically Complex	8	0.419	4.72	< .001	58	0.277	0.398

NOTE. – The combined effect size (r), combined Z, p, and fail-safe N are based on listed qualifying cases. Average percentage recall scores are based on qualifying cases from 216 studies using recall as a dependent measure.