INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book. These are also available as one exposure on a standard 35mm slide or as a 17" x 23" black and white photographic print for an additional charge.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700  800/521-0600
Issues in temporal coding

Pugh, Kenneth Richard, Ph.D.

The Ohio State University, 1989
ISSUES IN TEMPORAL CODING

DISSERTATION

Presented in Partial Fulfillment of the Requirements
for the Degree Doctor of Philosophy in the Graduate
School of the Ohio State University

By
Kenneth Richard Pugh, B.S., M.A.

The Ohio State University
1989

Dissertation Committee:
N.F. Johnson
L.E. Krueger
R.A. Fox

Approved by:
Advisor
Department of Psychology
To Jayshree and Priya
I gratefully acknowledge the contributions of Drs. Lester Krueger, and Robert Fox for their contributions to the content of this document. Special thanks are due my Advisor, Dr. Neal F. Johnson, whose guidance over the entire course of my graduate studies, as well as during the dissertation stage, has been indispensible. Working with him was both enlightening and enjoyable. I am grateful to my parents, Thomas and Vivienne Pugh for their loving support and confidence in me. Finally, I would like to acknowledge the person whose love, encouragement, and efforts, helped make all of this possible. A special debt of gratitude is owed to my wife Jayshree Pugh.
VITA


1982. B.S., New York Institute of Technology

1983-1985. Research Assistant, Sensory Biophysics, The Ohio State University

1985-1987. Teaching Associate, Department of Psychology, The Ohio State University


1988-1989. Instructor, Department of Psychology Denison University, Granville, Oh

PUBLICATIONS


FIELDS OF STUDY

Major Field: Experimental Psychology

Studies in: Language, Memory, and Cognition. Dr. Neal F. Johnson
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEDICATION</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>VITA</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
</tbody>
</table>

Chapter

I. INTRODUCTION ........................................ 1

Experimental Paradigms in Temporal Coding Research ........ 4

II. MODELS OF TEMPORAL CODING ..................... 11

Strength Theories .................................. 11
Time-Tag Theories ................................ 15
Contextual Coding Theories ....................... 19
Associative Coding Theories ..................... 23

III. A MODEL OF TEMPORAL CODING AND TEMPORAL JUDGMENT PERFORMANCE ........ 49

Assumption 1 ....................................... 50
Assumption 2 ....................................... 51
Assumption 3 ....................................... 61
Assumption 4 ....................................... 63
The Automacity Issue .............................. 67
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV. EXPERIMENTS 1-3: THE AUTOMATICITY ISSUE...</td>
<td>75</td>
</tr>
<tr>
<td>Experiment 1</td>
<td>77</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>87</td>
</tr>
<tr>
<td>Experiment 3</td>
<td>105</td>
</tr>
<tr>
<td>Summary of Experiments 1-3</td>
<td>110</td>
</tr>
<tr>
<td>V. EXPERIMENTS 4-5: THE STRENGTH BIAS ISSUE..</td>
<td>113</td>
</tr>
<tr>
<td>Experiment 4</td>
<td>116</td>
</tr>
<tr>
<td>Experiment 5</td>
<td>121</td>
</tr>
<tr>
<td>Summary of Strength Experiments</td>
<td>125</td>
</tr>
<tr>
<td>VI. CONCLUSIONS</td>
<td>127</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>137</td>
</tr>
</tbody>
</table>

APPENDICES

A. Mean Lag Values for Test Pairs from Experiment 4 147
B. Word and CVC Lists from Experiment 1 148
C. Test-pairs from Experiment 4 149
D. Associates used in Experiment 5 150
### LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Trial 2 mean Z' scores from Experiment 1...</td>
<td>83</td>
</tr>
<tr>
<td>2. Trial 2 mean Z' scores from Experiment 2...</td>
<td>91</td>
</tr>
<tr>
<td>3. Protocol data from Experiment 2.................................</td>
<td>92</td>
</tr>
<tr>
<td>4. Data from Experiment 3...........................................</td>
<td>108</td>
</tr>
<tr>
<td>5. Mean percent correct from Experiment 4.....</td>
<td>119</td>
</tr>
<tr>
<td>6. Mean deviation scores for priming conditions in Experiment 5.....</td>
<td>125</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURES</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Comparison of Chaining and Chunking Models</td>
<td>42</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

The ability to remember when an event occurred relative to other events has been interpreted as indicating that the encoding processes for that event must include the representation of temporal (or temporally relevant) information, and this temporal coding is considered to be one of the defining features of episodic memory (Underwood, 1969, 1977; Tulving, 1972; Michon & Jackson, 1984). Various theoretical accounts of temporal coding processes have been proposed, including time-tagging theories, contextual association theories, and strength theories.

The primary differences among these classes of theories concerns the specific type or types of information that constitute the temporal code (operationally defined as that information that mediates a temporal judgment). A secondary issue that differentiates some of these models is the question of whether temporal coding is considered an automatic or a controlled-process (Hasher & Zacks, 1979; Tzeng & Cotton, 1980). Obviously, hypotheses about the type of information that functions as a temporal code, and the question of whether such information is encoded automatically, are not orthogonal issues, and the current set of experiments are designed to shed light on both of these concerns.
In the first three experiments the question of whether temporal order information is encoded automatically, or whether the process is strategically mediated, will be addressed. Specifically, the failure of previous research to find evidence of intentionality effects will be explored. Subjects who are pre-informed before studying to-be-remembered materials that memory for temporal order will be tested perform no better than subjects not so informed (Toglia & Kimble, 1975; Tzeng, 1976). This failure to find expectancy effects has been interpreted as suggesting that the encoding of temporal information is an automatic process (Hasher & Zacks, 1979); one that operates outside of attentionally-controlled strategic processes (or at least is not materially affected by them).

Since the failure to find intentionality effects has really been the primary piece of evidence for the argument that "top-down" processes do not affect temporal coding, and since this seems at odds with most other memory phenomena, where subjects encoding strategies have a direct influence on memory performance, it becomes important to examine the issue. It is entirely possible that temporal coding is to some degree strategically controlled, but that the intentional/incidental manipulation simply fails to induce significant strategic shifts (Sanders, Gonzalez, Murphy, Liddle, & Vitina 1988; Naveh-Benjamin, 1987; Eyesenck, 1982).

Also, if it is determined that strategies can facilitate (or inhibit) temporal coding, an understanding of which strategies facilitate, and which inhibit, might allow inferences to be made
regarding what type(s) of information mediate temporal judgments (i.e., what constitutes the temporal code). Accordingly, an attempt will be made to identify relevant classes of strategies, and to determine whether they differentially affect temporal coding. This is especially important, since even after many years of experimental investigation there are many models, with quite diverse assumptions, that can account for performance about equally well, and there is no real consensus regarding the nature of the temporal code (a general theory will be proposed and tested here).

A second set of experiments will examine the decision processes involved in making temporal judgments. Specifically, they will explore the question of whether the temporal judgment processes, when more direct temporal information is unavailable, show systematic response biases based on the "strength" or availability of memories (Hinrichs, 1970; Brown, Rips, & Shevell, 1985). Such a notion, that memory-trace strength and perceived recency are positively correlated, dates back to the nineteenth century (Lipps, 1883; James, 1908; Michon & Jackson, 1984), but empirical support for this notion has been rare.

In these experiments certain manipulations are used that should, in principle, introduce response bias into temporal judgments by making items differentially retrievable. If, in fact, temporal judgments are biased by memory-trace strength, then any model that would account for temporal judgment performance would, of course, need to consider this bias in order to provide an adequate account of
performance. The conditions under which this bias is most likely to influence performance would need to be identified.

In the following section experimental paradigms that have been utilized to study temporal coding will be described. After that, prominent models will be described in some detail, and from this a general theory will be proposed. This will be followed by a discussion of the data relevant to the automaticity issue. Finally, the experiments conducted to examine these issues will be reported.

Experimental Paradigms in Temporal Coding Research

The most commonly used tasks in this field of research are: Serial-position judgments, lag judgments, and recency judgments. Each of these tasks is designed to measure memory for temporal order. The implicit assumption underlying the use of these tasks is that information that specifies when an event occurred in time will also allow judgments to be made about when the event occurred in relation to other events. Following this, a brief discussion of tasks used to study memory for naturally occurring events (events that occurred outside of the laboratory, and are part of the subjects autobiographical memory) will be presented. (It should be noted that the topic here is restricted primarily to memory for temporal information about essentially unrelated events. Research into memory for sequences with meaningful ordered relations (e.g., music, text) will not be included in this work).
Serial-Position Judgments

In this task, after studying a list of stimulus materials (e.g., words, pictures, etc.), subjects are shown an item from the list and asked to estimate its position. In some cases subjects are asked to give an exact numerical estimate, in others less precise responses are required (i.e., subjects might estimate which quarter, or which eighth, of the list an item occupied).

The most important finding is that when judged position is plotted against actual position using measures such as correlation scores, hits, or deviation scores, the data indicate that subjects consistently perform above chance, even with difficult materials and fairly long lists (e.g., 100 items). A regression towards the mean is also commonly observed, wherein early items tend to be judged as having occurred later in the list, and later items are judged as having occurred earlier. This has been found in both list-learning paradigms and in long-term dating of natural events, and probably reflects a general response bias (Bobko, Schiffman, Castino, & Chiapetta, 1977; Ferguson, & Martin, 1983). Another often reported finding is a tendency for subjects to be most accurate on the first few items from the list (primacy effect) (Tzeng, 1976).

Lag-Judgment Tasks

Two versions of this task have been commonly employed: paired-lag judgments, and absolute judgment of recency. In the paired-lag judgment task the subject is shown two items from a study list, and must estimate how many items intervened between them in that list.
Performance is usually at chance levels, with little improvement over trials (Underwood & Malmi, 1978; but see Proctor & Ambler, 1976). One exception to this pattern occurs when lag judgments are made for two items that are in some way related (e.g., two words from the same semantic category), or repetitions of the same items (Hintzman & Block, 1973; Hintzman, Block & Summers, 1975, Estes, 1985). Under these conditions performance is considerably more accurate than with unrelated items, and this has suggested to some researchers that these related items are rehearsed together and that this, in some (unspecified) way, establishes the temporal distance between the two.

In the absolute-recency judgment task, where subjects must estimate the lag between a test item and the end of the list (i.e., how far back in the list a test item occurred) the critical finding is that accuracy decreases with increased lag (Wells, 1974; Lockhart, 1969). Also, subjects tend to be more accurate on items from the end of the list (recency effect) (Wells, 1974).

In principle, this task is quite similar to the serial-position judgment task (a judgment about the position of a single target item is required in both tasks), yet the latter task tends to produce primacy effects, while the former produces recency effects. This difference probably reflects to some degree differences in memory search strategies used by subjects. Absolute judgment of recency instructions may encourage a backward scan which would tend to keep items from the end of the list more active in memory than items from the beginning. Serial-position judgment instructions, on the other
hand, may encourage forward scanning, and this in turn would keep items from the beginning of the list most active in memory. There is in fact evidence that indicates that, at least with short lists, recency instructions do encourage backward scanning (Hacker, 1980; Muter, 1979).

**Relative Judgment of Recency**

In this task, also known as paired-recency judgment, two test items are presented, and the subjects are asked to judge which item occurred more recently than the other one. This task would seem to demand the least precise temporal information of those discussed thus far.

Three common effects are reported in this task. First, the greater the lag between the two items the more accurate are the judgments (Michon & Jackson, 1984; Yntema & Trask, 1963). Second, for any two target items separated by the same lag, the greater the lag between the more recent item and the test stage, the less accurate the judgments tend to be (Yntema & Trask, 1963; Estes, 1986). The negative effect of an increasing retention interval in this task, as well as the absolute-recency judgment task, would seem to reflect forgetting of item and/or order information over time. Third, as with lag judgments, accuracy is greater on pairs of items that are semantically related (Tzeng & Cotton, 1980; Winograd & Soloway, 1985).

**Comparison of these three types of tasks.** These tasks differ in a number of ways, but the most obvious difference concerns the relative precision demanded. For example, the paired-recency task requires
only that the subject retrieve enough information to determine which item occurred later in the list, but the serial-position, lag, and absolute-recency tasks demand more precise estimates of a given item's location. Therefore, it is entirely possible that different types of memorial information can mediate these different judgments. However, there is evidence provided by transfer of training studies, in which training with one type of task facilitates subsequent performance on others (Underwood & Malmi 1978), that suggests that the same type or types of information may be mediating the different judgments (but see Michon & Jackson (1984) for contrary arguments). The question, of course, is what type of information does mediate the tasks, or stated differently, what is the nature of the temporal attribute(s).

**Temporal judgments for natural events**

The relative judgment of recency (paired-recency) task has been used in assessing temporal coding in memory for natural events (i.e., events that occur outside of the psychological laboratory and that are part of the subject's autobiographical memory) (Brown, Rips, & Shevell, 1985). In this type of task certain evidence has been obtained which suggests that subjects tend to be biased by the amount of information that they can recall, judging more vivid memories as more recent; an apparent strength bias (Brown, Rips, & Shevell, 1985; but see Huttenlocher, Hedges & Prohaska, 1988 for alternative interpretations of this data).

The standard distance effect has also been obtained (Fuhrman & Wyer, 1987), wherein the further apart in time two events occurred the
more quickly and accurately their relative order is judged. Further, events from different temporally defined episodes (e.g., high school vs. college) are more quickly and accurately judged than items from the same episode, suggesting that subjects make use of higher-level segmentations in making these judgments (Fuhrman & Wyer, 1987).

Another paradigm that has been used with natural events involves having subjects keep a diary of important events over a long period of time, and subsequently asking them to assign dates to each of the items (the experimenter is thus privy to the true dates which the subject originally recorded). A commonly reported effect is the so-called "telescoping error", where for less recent items especially, subjects tend to date the event as having occurred later than it actually did (Thompson, Skowronski, & Lee, 1988). The tendency to overestimate recency has been reported in list learning paradigms as well (see above), and may be indicative of more general response biases (however, see Huttenlocher et al. (1988) for an alternative interpretation).

There is no reason to assume that the information that specifies temporal position with unrelated stimulus lists in the psychology laboratory and the information that mediates recency or dating judgments for natural events is the same, and in fact, it is unlikely that it is. In autobiographical memory many more temporal "markers", such as key events, or literal calendar information (memory for specific dates), are likely to be used in making judgments. Such information would, of course, be unavailable when making judgments.
about the temporal order of a list of stimulus items presented in a laboratory task. However, we may observe parallels at the level of the judgment processes themselves and in the types of response biases they exhibit (e.g., strength bias, use of anchor points in making judgments). For example, in both types of paradigm subjects may be using complex inferential strategies to reconstruct order or assign a position (or date) to an event, and while the specific information that is utilized may differ considerably, the processes themselves may not (Brown, Shevell, & Rips, 1986). For this reason it is useful to keep track of similarities in findings obtained from these different paradigms.
CHAPTER II
MODELS OF TEMPORAL CODING

Four distinctive classes of temporal coding models have been proposed in this literature: Strength theories, time-tag theories, contextual coding theories, and associative coding theories (Underwood, 1977). The type of information that would mediate a temporal judgment is quite different in each of these formulations. In strength theory, temporal position is derived from the magnitude or strength of the memory trace for the item, which is assumed to decrease monotonically with time. In time-tag theory the memory representation of the event includes an attribute that marks directly its position in time. In contextual coding theory various internal and external contextual associations are thought to preserve temporal information. In associative coding theory, associations between input items formed during the study stage, can be used to derive an estimate of a given item's position in the series. Each of these theories will be discussed in detail in this section.

Strength Theories

In this class of models it is assumed that temporal judgments are made by using the strength or magnitude of the memory trace for an
event to estimate how recently the event occurred. The strength is thought to decrease monotonically with time, and the subject, in some way aware of this fact, will believe that the "weaker" (or less vivid) the memory-trace, the less recently the event must have occurred (Hinrichs, 1970; Wickelgren, 1974).

The simplest model of this type is known as the single-trace strength theory (Hinrichs, 1970; Flexser & Bower, 1974; Hintzman & Block, 1971). In this model, it is assumed that for any given stimulus item the subject will have a single memory representation, which does not contain any information about when the event occurred, or how many times it was repeated (frequency information). However, every time the item is encountered the strength value of the memory trace is increased (labeled single-trace summation). The model was originally proposed to account for the fact that the likelihood of both recall and recognition increases when an item is presented more than once. Quantitative versions of this type of model have been proposed, and these can quite adequately account for the various lag effects that were described in the section on experimental paradigms (Hinrichs, 1970; Wickelgren, 1974; Bower, 1972).

Both temporal judgments and frequency judgments are made by translating the strength value of the memory trace into either type of estimate. The fact that neither type of information is represented directly leads to an interesting prediction regarding the relationship between repetition and judged recency. If repetitions increment the strength value, and if this strength is used to estimate recency, then
repeated items should be judged to have occurred more recently than they actually did occur (recency should be overestimated). Morton (1968) tested this prediction in a paired-recency task, using digits as stimuli, and found that repetition of the less recent item resulted in decreased accuracy, while repetition of the more recent item increased accuracy. That result is consistent with the predictions derived from single-trace strength theory. Fozard and Yntema (1966) reported the same effects with pictures as stimulus items.

Peterson and his colleagues (Peterson, 1967; Peterson, Johnson, & Coatney, 1969) failed to obtain this effect when repetitions were spaced, but did obtain it when repetitions were massed. However, since spaced repetitions result in a greater probability of recall than do massed repetitions (Crowder, 1976), this failure to demonstrate the strength bias in the spaced repetition condition is problematical for the simple strength model. Further, although Flexser and Bower (1974) did obtain evidence that repetition of the less recent item led to greater error in paired-recency judgments as predicted, they also found that repeating the more recent item in the test pair resulted in greater error when subjects were asked to judge which item occurred less recently (primacy task). This reversal is not predicted from the strength model. The data, then, on the repetition effect are equivocal.

Also, the notion of single trace summation has not fared well, in general. There is considerable data supportive of the notion that each repetition of an item results in at least partially independent
representations, and this is not consistent with trace summation accounts (Hintzman, Block, & Summers, 1973; Hintzman, 1976). The central assumptions of strength models, therefore have not received unequivocal support in the literature.

There is also considerable evidence that subjects often have available to them more direct information when making temporal judgments (see below), and therefore it is clear that there are many conditions under which strength models could not account for performance. Nonetheless, the real question concerns whether, under conditions where more direct information has been forgotten (or is not initially encoded), subjects will rely on perceived strength (or vividness of the memory) to estimate when the event occurred. While this expectation seems plausible to many researchers, in fact, at present there is little evidence that would support the hypothesis. Given that more direct temporal (or temporal relevant) information appears to be quite rapidly forgotten, it is important to determine whether subjects under these conditions will use strength-like cues or simply resort to guessing (and Experiments 4 and 5 address this issue).

In any event, a simple strength model would not be capable of handling many of the findings from this literature, and so it would appear necessary to look to either time-tag models, or associative coding models, to provide an adequate account of temporal order learning, in general.
Time Tag Theories

In this type of model it is assumed that any memory representation will include an attribute that marks the temporal location of the item or event. Most of the researchers who use the term do not specify just what this tag is, but assume it is invariant to situational or contextual differences. That is, it is assumed that the time tag is a fundamental memorial attribute, and one that, in some way, represents time of occurrence in direct fashion (Treisman, 1963; Flexser & Bower, 1974; Glenberg & Swanson, 1986). In each of the models to be discussed in this section an interval scale representation of temporal information is assumed. This means that not only is ordinal information encoded, but the temporal distance or interval between any two items is also maintained.

Treisman (1963) argued that the time-tag represents the current value on a kind of biological clock. That is, when the event is encoded the representation will include the current clock value (see Bindra, 1976, Kristofferson, 1977, for similar ideas). This clock would presumably be set to a zero value at the beginning of a distinctive episode, and would be reset to zero at the end of the interval (Kristofferson, 1977). Models of this sort have primarily been applied to studies of memory for temporal duration.

Murdock (1976) proposed that the memory representation for a series of events is analogous to a tape recorder, with each event registering at a different point on the moving "tape". That would make position on the tape correlated with position in time. This model is
conceptually similar to the proposal that subjects convert temporal position into a spatial analog format (imaginal representation). That is, position in time is represented in memory as position in a spatial image. Again, this type of representation could preserve not only order information, but also interval or distance between events (Michon & Jackson, 1984). The notion that imaginal representations can preserve interval scale information has received some empirical support in the literature on linear-order learning (Huttenlocher, 1969; Moyer & Bayer, 1976; Shepard, Kilpatric, & Cunningham, 1975).

Several other researchers have proposed time-tag theories to account for long-term recency effects in recall tasks (Glenberg & Swanson, 1986; Glenberg & Fernandez, 1988; Bjork & Whitten, 1974; Flexser & Bower, 1974; Gardiner & Gregg, 1979). Each of these models assume that time of presentation is encoded on an interval level scale, but none of these models specifies exactly what type of information is contained in the time tag.

In each of these types of models, which appear to be quite different on the surface, there is an assumption that true temporal information is encoded, and encoded in a way that is, in some sense, independent of other events occurring in the same episode. This contrasts sharply with the assumption found in associative encoding theories (see below), that an item's position relative to other events is encoded through associative mechanisms and that this serves as the cue to its temporal position (the emphasis there is on relational coding processes).
A model that assumes an interval level representation of temporal information makes the following prediction in a temporal judgment task. The further apart in time the items are spaced the more accurately the position of any given target could be estimated (Glenberg & Fernandez, 1988; Greene & Crowder, 1988). This follows from the idea that the time-tags for items that are adjacent in the list will be more dissimilar the greater is the temporal interval between them. For example, the values of the time tags for items 6 and 7 in a ten-item list will be similar if they are presented with a short inter-stimulus-interval (ISI), and dissimilar if the ISI is increased. If these items were presented as a test-pair in a paired-recency judgment the more dissimilar the time-tags the more accurate the judgment should be (Glenberg & Fernandez, 1988). Therefore increasing the ISI's should result in more accurate temporal judgment performance in general.

Glenberg and Fernandez (1988) tested this prediction using a paired-recency judgment task. Presentation of items was either massed or spaced (separated by distractor filled ISI's). While accuracy for pairs of items from the beginning or end of the list did increase with longer ISI's, this advantage did not obtain for mid-list items. The time-tag account predicts a spacing advantage across all input positions, therefore these results do not unequivocally favor the time-tag theory.

Three other studies have manipulated ISIs to determine whether increased temporal spacing will facilitate temporal judgments, and
have failed to obtain a spacing advantage. Guenther and Linton (1975) varied either the temporal lag, or number of intervening items, between two test items in a paired-recency judgment in order to determine which factor was responsible for the standard distance effect (increased accuracy with increased lag). They found that accuracy increased only as the number of intervening items was increased, and not as a function of the temporal interval between the items. Therefore, these data fail to support the prediction of interval level time-tag theories.

Greene and Crowder (1988), using a serial-position judgment task with short lists, found that performance was more accurate in a massed presentation condition than in one with filled ISIs. Also, Shiffrin and Cook (1978) reported similar effects using a short-term memory paradigm and serial recall task. In both of these studies, increased spacing actually inhibited performance in a temporal order task. Although spacing does facilitate recall (Bjork & Whitten, 1974), presumably by producing more distinctive encoding for items, it appears to inhibit the type of encoding useful in temporal judgment tasks.

In short, the critical prediction from interval level time tagging theories is that increasing the temporal distance between items will lead to more distinctive codes for items by decreasing the similarity of the time tags, and this should result in more accurate temporal judgment performance (Glenberg & Fernandez, 1988, Greene, & Crowder, 1988). The data to this issue do not appear to be inconsistent with
this prediction. In general, increasing the ISI's between items decreases accuracy on tasks that require temporal order information. There is other evidence that will be presented in the next section that would suggest that subjects rely on various kinds of associative information when making temporal judgments, and these findings also seriously challenge the idea that true time of presentation is encoded in memory. Therefore, at present at least, there is no evidence that true temporal information is encoded in the memory representation for an event.

**Contextual Coding Theories**

Each of the models to be discussed in this section posits at least one type of contextual information that serves as a temporal code. In this way each one contrasts with the strength theory account, where temporal information is not directly represented in memory, and the time tag theories where true interval level temporal information is represented. Given that these latter two classes of models have not received much empirical support, it is probably the case that the key to temporal order learning lies in the kinds of associations that subjects form when studying to-be-remembered items. A list of these hypothesized sources of contextual information includes associations between target stimulus events and: 1) extra-list context, and 2) internal cognitive processes that change systemically with time. These ideas, and relevant empirical findings will be discussed in this section.
Extra-List Context

The idea that to-be-remembered items in a list-learning paradigm get associated with objects or events that are contiguous with them (occupy working memory at the same time) is richly supported in the memory literature (Tulving, 1983). That such extra-list context might specify when in the list a given item occurred, would require that this contextual information be inherently ordinally meaningful; in other words, that it preserves ordered relations (Underwood, 1977). For example, if along with the to-be-remembered list items, other events that are well-structured, were presented as a background context, they might get associated with the target items, and help preserve temporal order information.

Evidence that extra-list context influenced temporal judgment performance was reported by Guenther and Linton (1975), and Michon and Jackson (1984). Guenther and Linton found that when a well structured chronologically organized story was presented simultaneously with a series of unrelated randomly ordered pictures that subjects were to remember (also unrelated to story), paired-recency judgments for the pictures were more accurate than in a control condition, where the pictures were presented alone.

Michon and Jackson (1984) essentially replicated this effect, but they also found that there are limiting conditions. Specifically, when the to-be-remembered items (in their case words lists) are associatively related, or otherwise coherently organized, extra-list context seems to be less important. That would suggest that perhaps
subjects allocate less attention to the background context when the critical items encourage elaborative rehearsal strategies. In any event, it does appear that events that co-occur with the stimulus list can become associated with the items, and serve as temporal codes under certain conditions. However, under conditions where the study list is presented in isolation other mechanisms are required to account for temporal order learning. In any event, these studies demonstrate that extra-list context can facilitate the process of learning the order of stimulus items.

**Internal Context**

If cognitive states change in some non-random fashion across the study-trial, and if these states become part of the representations for the stimulus items, then they might mediate judgments of temporal order. For example, changes in mood or attentional focus across the list, or changes in cognitive processes across the list as the subject attempts to cope with the demands of the memory task, could potentially get encoded with the to-be-remembered items, and facilitate temporal judgments.

Research on state-dependent memory (Overton, 1972; Spear, 1977; Bower, 1981) has indicated that a subject's internal state at encoding may become part of the memory for the stimulus event, since recall is usually better when, for example, mood at retrieval matches mood at encoding (Bower, 1981). This effect has also been observed with drug and alcohol manipulations (Overton, 1972; Eich, Weingartner, Stillman, & Gillin, 1975). With this in mind, it has been conjectured that if
these internal states change in some temporally relevant fashion across the trial, then recovery of this information might function as a temporal code. For example, Underwood (1977) discusses the possibility that at the beginning of the study list the subject is anxious, toward the middle bored, and near the end impatient, and if these moods get associated with the target items then temporally relevant information is indeed preserved. However, to date there has been no empirical verification of this notion, but a number of researchers appear to find the notion appealing (Bower, 1974; Hintzman, Block & Summers, 1973; Underwood, 1977). In short, while general contextual factors might facilitate temporal coding, it is quite difficult to document these effects, since they reflect idiosyncratic processes that are quite likely to differ from subject to subject.

Process context reflects the cognitive states that result from the cognitive operations that the subject is employing in the experimental task (Underwood, 1977; Block, 1982). Unlike mood, this task specific internal context can be controlled by the experimenter by manipulations of orienting task. Block assumes that the ways that one thinks about, and processes events, becomes part of the memory for that event. For example, decisions made, associations with existing knowledge, and any other cognitive processes that might be active would become part of the representation of an episodic memory. This idea falls quite naturally out of the information-processing framework, where it is assumed that people are not passive recorders,
but rather, are active encoders (Lachman, Lachman & Butterfield, 1979). If that is coupled with the assumption that process context is stored in memory, and if it changes over the trial in some way that is temporally meaningful, then it might function as a temporal code.

Underwood (1977) found that inducing subjects to engage in different orienting tasks for different lists improved list discrimination judgments, compared to conditions where the same orienting task was used across lists. Block (1982) reported similar findings using a two list discrimination task. This implies that the process context manipulation did indeed improve temporal discrimination. Apparently, when subjects were shown a target item, they could remember how it was encoded, and could use this information to remember in which list the target item occurred.

**Associative Coding Theories**

Just as various kinds of internal and external contextual information may be active in working memory when subjects are studying a to-be-remembered item, and may get associated with it, it is also the case that other items from the list may get rehearsed with the item, and might help preserve the item’s temporal position. The simplest account of this sort is contained in serial-chaining models (Shiffrin & Cook, 1978; Lewandowsky & Murdock, 1989). In chaining models, directional associations are formed between adjacent items, and from these associative processes the temporal order of the list would be preserved. If a temporal judgment for an item or pair of
items is required, then the subject would engage in implicit serial recall, traversing through the chain until the target(s) is found (Shiffrin & Cook, 1978).

It has also been suggested that groups of items that are temporally contiguous might get rehearsed together, resulting in hierarchically organized chunks, or distinctive episodes (Johnson, 1972; Estes, 1986). Since there would be fewer chunks than actual list items, this should have the effect of creating a shorter list at the chunk level (e.g., for a twelve-item list, chunking into groups of three would result in four chunks, and if the order of these chunks is encoded, as well as the order of the items within a chunk, then the temporal order of the list is preserved in an efficient schema). Chunking has been shown to facilitate serial recall (Lee & Estes, 1981; Nairne, 1988), and would presumably facilitate temporal position judgments for single items or pairs of items as well, by providing higher order segmentation that would allow an item's relative position in the list to be derived.

It will be shown below that chaining and chunking models make different predictions in temporal judgment tasks, and they can be contrasted directly. However, both assume that temporal order learning is dependent on localized rehearsal patterns; that is, a given item must be linked in some fashion with other items that are temporally contiguous with it (in the chaining model it would be the immediately preceding and following items, and in the chunking model groups of adjacent items in some hierarchically-organized fashion).
On these accounts, rehearsal patterns that encourage associations between items that are not temporally contiguous, as when an item triggers the retrieval of an item from earlier in the list that shares some pre-existing associative relationship, would interfere with the processing involved in associating the item with its temporal neighbors, and should inhibit temporal order learning.

In contrast with the chaining and chunking models, some researchers have suggested that displacing rehearsals in this fashion will actually lead to better temporal order learning, because in rehearsing non-adjacent pairs, both the order and the distance, or lag, will get encoded, and this might establish a good temporal-order representation (Tzeng & Cotton, 1980; Winograd & Soloway, 1985; Zacks, Hasher, Alba, Sanft, & Rose, 1984). However, like the chaining and chunking models, the displaced rehearsal hypothesis assumes that temporal order learning is dependent on relational coding among input items, and no representation of true interval-level temporal information is assumed. In that way, each of these models contrasts with time-tag theory.

In this section some data that has supported the idea that relational coding among input items is critical for temporal coding will be presented. Following that, the evidence that suggested a role for displaced rehearsal will be described. Finally, certain formal serial-learning models that propose either chaining or chunking accounts will be described in some detail.
Evidence of Associative Coding

One finding that would appear to support the assumption that associations formed between a given item, and other items from the list, will result in more accurate temporal coding for the item, was reported by Tzeng, Lee, and Wetzel (1979). In their experiment the directed-forgetting paradigm was used (Bjork, 1972), wherein immediately after the presentation of an item the subject is instructed either to remember it (to-be-remembered item TBR) or to forget it (to-be-forgotten item TBF). Subjects were expecting a free-recall task, and after completing it they were given a surprise serial-position judgment task. The TBR items were recalled more accurately, and serial-position judgments for these items also were more accurate than for TBF items. That is consistent with the notion that maintaining an item in the memory or rehearsal set will improve temporal coding (presumably through associative coding with other items). However, it is also possible that the advantage for TBR items in the serial-position judgment task was simply the consequence of the fact that they were better remembered than TBF items, and so the experiment confounded these issues (Logan, 1988).

In a second experiment the authors once again used the directed-forgetting paradigm, but followed the recall task with an unexpected paired-recency judgment task. The test pairs consisted of all combinations of to-be-remembered (R) and to-be-forgotten (F) words. Performance was most accurate with R-R pairs, second best with R-F pairs, and equivalent for both F-R and F-F pairs.
The critical support for the associative coding idea is seen in the fact that performance was better on R-F pairs (where the more recent item is TBF) than on F-R pairs (where the more recent item is TBR). A simple forgetting account of the advantage for TBR items would predict that the item that is remembered will be judged as more recent, and on this account F-R should be more accurate than R-F pairs, but just the opposite pattern of results was obtained. Tzeng et al. (1979) argued that because the earlier item is maintained in the rehearsal set it will get directionally associated with the later TBF item, even though the latter item is subsequently dropped from the rehearsal set. However, in the case where the earlier item is TBF it is unlikely that its position relative to the later TBR item will be encoded (they will not be associatively rehearsed), and hence performance on such a pair is less accurate. In summary, instructions to retain an item result in more accurate temporal judgments for the item, presumably by encouraging associations to be formed between this item and other list items.

It should be noted that Tzeng et al. (1979) consider these associations that function as temporal codes to be automatically encoded (Tzeng & Cotton, 1980). By automaticity, they simply mean that presence in the working memory rehearsal set is sufficient to produce the appropriate associations, and that subjects intentions and strategies will not add (or subtract) anything to this process. In support of this claim they found no difference between incidental and intentional conditions in either serial- position judgments or paired
recency-judgments (Tzeng, 1976; Tzeng, Lee, & Wetzel, 1979; Toglia & Kimble, 1976; Hasher & Zacks, 1979). (The implications of this claim will be discussed in a subsequent section.)

**Displaced Rehearsal Theory**

Building on the associative coding idea, Tzeng and Cotton (1980) examined the role of displaced rehearsal in temporal coding. They reasoned that whatever co-occupies working memory with an item that is being currently studied, will get associated with that item, and this association can serve as a temporal code. It is likely that while studying a given item the past few items still would be in memory (perhaps up to the span limits of working memory), along with internal and external extra-list context (see above). However, along with these items any earlier items that are associatively or semantically related to the current item, might be automatically retrieved from LTM and also would co-occupy working memory. Because the earlier item is retrieved from memory, and the current item is being perceived, Tzeng and Cotton postulate that the relative order of the two items will be automatically established. They further argued that the lag between these items will also get encoded, and cites data from Hintzman, Summers & Block (1975) that supports this claim. If a sufficient number of displaced rehearsals are performed, then a representation of the temporal order of the list will be established. In fact, they go so far as to claim that displaced rehearsal is the central mechanism in temporal-order learning (this, of course, conflicts with the idea contained in chunking and chaining models, that localized rehearsal
patterns are crucial for temporal-order learning).

The fact that related items tend to get rehearsed together, even when they occur at relatively long lags, was first established in a lag judgment experiment conducted by Hintzman, Summers and Block (1975). Subjects were instructed to estimate how many items had occurred between two target items. Three types of test pairs were used: 1) Repetition of the same item (e.g., beer-beer); 2) associatively related items (e.g., beer-vine); or 3) unrelated items (e.g., beer-goat). The correspondence between judged lag and true lag was greatest for the associatively related pairs, second for the repeated pairs, and almost non-existent for the unrelated pairs. They concluded that when an item triggers retrieval of an earlier one, the temporal relationship between them is likely to be established (both the distance or lag, and the relative order).

Tzeng and Cotton (1980) confirmed the prediction that relative order information will be established between items that are likely to be rehearsed together using a paired-recency judgment task. In their first experiment half of the subjects studied a categorized word list, and the other half studied a list consisting of unrelated words. The categorized list consisted of 10 categories with five instances of each, presented in random order. After the study phase an unexpected paired-recency judgment task was given to subjects. Test pairs were either intracategory (beer-vine) or intercategory (beer-goat) for the categorized list. The same intracategory and intercategory assignments were given to items from the same serial positions in the
unrelated list, but since there were no categories in this list, this manipulation just served as a control condition.

The results were consistent with the displaced rehearsal hypothesis. Subjects were most accurate for pairs of related items (intracategory) from the categorized list. Overall, however, there were no differences between the unrelated list and the categorized list. The conclusion drawn by the authors is that, in general, temporal information is established for items that are rehearsed together. However, the likelihood of co-rehearsal is greater when the words share some pre-existing associative relationship, as evidenced by the advantage for intracategory pairs from the categorized list, and that is true even when these words are not temporally contiguous.

Winograd and Soloway (1985) replicated this effect, and by instructing subjects to think of an associate for each word studied and mark it down, they found that when the associate was an item from earlier in the list, paired-recency judgments for these pairs were more accurate than for pairs not rehearsed together. That result, of course, supports quite convincingly the claim that displaced rehearsal establishes relative temporal order information.

The unanswered question in these studies is whether displaced rehearsal patterns, that do facilitate paired-recency judgments for items rehearsed together, actually facilitate or interfere with the learning of temporal order in general. For example, if the task required serial-position judgments for single target items would
displaced rehearsal patterns produce equivalent performance to localized rehearsal patterns? Displaced rehearsal theory as espoused by Tzeng (Tzeng & Cotton, 1980) argues that displaced rehearsal patterns would produce superior performance.

The reasoning is as follows: All associations are formed automatically (without requiring attentional effort). Therefore items that are contiguous (are currently in working memory), will get associated, and therefore localized rehearsal is insured. Further, conditions that induce displaced rehearsals, such as would be the case when an earlier item is an associate of the current item, will establish the temporal relationship (order and lag) between the items in working memory and the items that were contiguous with the earlier associate. In this way an ordered representation is established. On the other hand, models that assume that localized rehearsal patterns are critical for temporal order learning, and do not assume automaticity of associative coding, would predict that the displaced rehearsals would interfere with order learning, since this would interfere with chunking or chaining processes.

The data relevant to this question would seem to support the idea that displaced rehearsals interfere with temporal-order learning. Tzeng and Cotton (1980) found that while performance on intracategory pairs was more accurate than on intercategory pairs from the categorized list suggesting that displaced rehearsals were occurring, performance on the categorized list was not superior to that on the unrelated control list. Thus, displaced rehearsals did not facilitate
order learning in general. However, it is also the case that it did not seem to interfere with order learning since overall performance on the two lists did not differ, and this can be taken as partial support of Tzeng’s automaticity hypothesis. On the localized rehearsal hypothesis performance should have been superior on the control list, since displaced rehearsal patterns should be less likely with unrelated items.

However, Pugh (1986) reasoned that a true shift to displaced rehearsal strategies (clustering items from the same category that were not temporally adjacent in the list) might not actually have occurred in the Tzeng and Cotton (1980) study. In that study, with one study trial and conditions where subjects were not pre-informed about the categorical nature of the list, clustering might have been quite limited. He hypothesized that if more than one study trial were given, subjects studying the categorized list might shift to true clustering strategies, and here the adverse effect of displaced rehearsal patterns might be seen. Consistent with this, he found that with four study trials prior to the paired-recency task, performance on the control list was superior to performance on the categorized list. This would seem to imply that patterns of displaced rehearsal can, in fact, interfere with temporal-order learning. This finding would appear to be inconsistent not only with displaced rehearsal theory, but with the automaticity hypothesis in general (see below).

In any event, in the formal serial-learning models discussed next, localized, as opposed to displaced, rehearsal patterns are necessary
conditions for temporal-order learning, and these ideas can be
contrasted empirically with displaced rehearsal theory.

**Formal Serial-Order Learning Models**

The idea that relational coding among input items can preserve
temporal order information, has received a more formal expression in a
number of models generated to account for serial learning, such as
1981, Estes, 1986), or Shiffrin (Shiffrin & Cook, 1978). In each of
these models, which do differ on a number of detailed points (for
example, whether simple chaining or chunking processes are included),
an attempt is made to account for how subjects represent order
information, and how they use this information to recall a series of
items in the same order as they were presented (serial recall). Each
of these models assume that localized patterns of rehearsal are
crucial to order learning, and consequently they can be directly
contrasted with displaced rehearsal theory (Tzeng & Cotton, 1980;
Winograd & Soloway, 1985). Further, given the lack of empirical
support for true time tag theories and strength models, and given that
there are many experimental conditions where extra-list contextual
information is minimal, it might prove fruitful to turn to this class
of models for a reasonable account of temporal-order learning.
These models will be briefly examined, and the ways in which they
might be extended to the types of tasks that are being discussed here,
will be considered.
Johnson (1970). Johnson in his coding model (1970, 1972, 1978) postulates that when a series of items is presented to a subject there will be a tendency to group subsets of these items into chunks (operationally defined as subsets of items that tend to be recalled adjacently, and in an all-or-none manner). Each of these chunks in turn is represented by a unitary code which is logically distinct from the information that it represents (codes are viewed as opaque containers). These codes do, however, represent all the information in the chunk including item order. These codes, in turn, may be components of still higher level codes that represent, for example, the order of these chunk codes. The generative process during recall proceeds from the top of the chunk hierarchy down to the response elements. That is, the subject will first decode higher-level codes which may contain the ordered chunk codes, and only then decode the first chunk and produce the response items from this chunk in their appropriate order. That would be followed by decoding the second chunk, and so on.

A critical assumption regarding the current issues is that an item's serial position is represented by this hierarchical coding structure (i.e., its position within a chunk is represented in that chunk code, and its position within the series is further represented by the higher-order codes of which it is a part). Johnson argued that this positional information is represented within the codes by position tags; therefore an item's position in the list is represented by both within-chunk position tags, and the higher-order codes that
represent chunk order. He further proposes that given the evidence suggesting that the optimal chunk size is three items, it is quite possible that within-chunk position might be contained by codes corresponding in some way to the concepts initial-medial-terminal (note similar arguments from Drewnowski (1980) and Underwood (1977)).

Importantly, no item-to-item associations are assumed. In fact, the notion of hierarchically organized position coding, also evident in Estes model, was motivated by the failures of strictly associative models of serial-order learning (such as the serial-chaining hypothesis) to account for data from transfer tasks (Young, 1968; Johnson, 1970; Shiffrin & Cook, 1978; but see Levanowsky & Murdock, 1989), or to handle subjects ability to rapidly produce sequences of responses (Lashley, 1951; Johnson, 1970).

Originally this model was proposed to account for performance in a serial-recall task, with top-down retrieval of codes through a process of decoding, and subsequent decoding into component information, and consequently it does not provide an account of recognition, in general, or temporal judgments in particular. However, if the position tags within the hierarchical coding scheme could be accessed in a bottom-up fashion when the stimulus item is presented in a temporal judgment task, then, of course, they might (and probably would) mediate a temporal judgment decision.

The results of one experiment conducted by Johnson (1978) would seem to support this idea. Subjects were instructed to memorize a string of nine letters, and chunking into groups of three was induced
by spatially separating the third and fourth, and sixth and seventh, items. After learning the sequence, the subjects were shown rows of nine dashes with a single letter appearing on one of the dashes. The subjects' task was to decide whether the letter was in the correct position. Reaction time was the critical dependent measure.

The results seem quite consistent with the notion of hierarchically-organized position coding in that the time taken to determine that an item was in an incorrect position decreased both with physical distance from true position, and with increasing numbers of discrepant order tags within the assumed hierarchy. It seems quite reasonable to adapt the model in this fashion to temporal-judgment performance, and as will be seen below Estes proposes a very similar model, and he does adapt his model to account for temporal judgments.

There are a number of ways in which this model of serial learning might provide insight into the cognitive processes that operate to establish higher-order organization for sequences of items or events that unfold in time, and it is these insights that might be incorporated in a general model of temporal coding. First, chunking effects, which are legion in list learning paradigms, have been largely ignored in those models which were designed to account for the specific process of making temporal judgments. That is strange, because chunking and hierarchically-organized segmentations provide an explicit manifestation of the rather vague notion of process context, in that this organization is the result of cognitive
processes that operate on input (Lee & Estes, 1981; Block, 1982; Underwood, 1977). That is, the cognitive processes of chunking and hierarchical coding serve to partition a series of items into functionally distinct segments, which in turn are ordered. When there are no external extra-list markers, it is quite probably these hierarchical structures that function as temporal context (Estes, 1986).

As more and more real world structure is represented in a sequence of events, the organizational processes would obviously differ in detail, but probably not fundamentally (e.g., chunking might be based on semantic relationships as well as temporal contiguity, and the "meaning" of the events would be part of the organizational structure or context). For example, most recent theorizing on temporal organization in autobiographical memory (memory for natural events that span a lifetime) assumes hierarchically-arranged temporal units which serve to place an event in temporal perspective (Robinson, 1986; Neisser, 1986; Huttenlocher, Hedges, & Prohaska, 1988). There is evidence from text comprehension and recall indicating hierarchical organization and top-down breadth-first retrieval processes (Kintsch & Van Dijke, 1978; Thorndyke & Yekovich, 1980). The key point is, that a formally stated model of this type, with its explicit encoding and retrieval assumptions, provides a framework that can be extended upward to situations with more contextual richness, and beyond the tasks for which it was designed to account.
Estes (1972). Estes model of the encoding of item and order information in short-term memory has undergone a number of revisions over the years (Estes, 1972, 1986; Lee & Estes, 1977, 1981). The most updated version is found in Lee and Estes (1981), and a discussion of how it might be applied to temporal-judgment tasks, such as serial-position or recency judgments, is found in Estes (1986). Like Johnson’s (1970) model, order information is represented by hierarchically-organized structures which preserve serial position information for each item, and importantly, no item-to-item associations are assumed. Specifically, each item is associated with hierarchically arranged control elements that function as temporal context for the item. One control element represents position within a segment, at a higher level the position within a list is represented, and at a still higher level the item is associated with control elements that specify the list or trial on which the item occurred.

Any other cognitive processing that is applied to the item also will get associated as context. However, associations between adjacent items occur at the level of control elements not directly between the items themselves, and this notion is built in for the same reasons as in the Johnson (1970) model (i.e., the failure of such models to account for failures in transfer tasks, etc.). The Estes model contains chunking organization within the hierarchy, but the details of the coding structures are not as clearly laid out as in the Johnson model.
This mathematical model of short-term memory organization makes its most unique contribution with an elaborate account of the forgetting of item and order information. Lee and Estes (1981) argued that sequences of items will be rehearsed intermittently, or reactivated after initial storage (this process need not be a conscious one), and there is some probability that a given item will be reactivated out of the original sequence. That perturbation will lead to a loss of position information for that item (proper control elements are lost). Thus, over time, due to reactivation, item-position information will be lost and this perturbation process is assumed to be random. Item information is only lost because the spatio-temporal context in which it was embedded is lost. In other words, loss of order information produces loss of item information (see Anderson & Bower, 1972 for similar ideas).

In applying this multi-level contextual-coding model to temporal-judgment performance, Estes (1986) assumes that presenting two items for a paired-recency judgment will make available all of the control elements associated with each item. With increasing distance between the items in the list it is more likely that higher-level control elements will be different, and that can account for increased accuracy with increased lag (distance effects). Similarly, since the perturbation of order information at multiple levels is more likely with longer retention intervals, the model also, in principle, can account for the decreased accuracy as the lag of the more recent item increases. (Note that the Johnson (1970) model would account for both
lag effects similarly, but while the structural assumptions might be more explicit in the Johnson model, the mechanics of forgetting are not worked out to the same degree as in the Estes (1986) model.) Of course, the problem for both models would be to specify the nature of the hierarchical structure when there are long lists of items with no experimenter provided segments to induce chunking schemes. Nonetheless, any sort of processing which produces higher-order segmentation, including elaborative strategic processes, could be included in these flexible models.

Shiffrin (1978). Shiffrin and Cook (1978) developed a model of short-term memory for item and order information that, unlike the previous models, does assume a serial-chaining hypothesis; that is there are direct associative links for adjacent items. The authors argue that time tags are not established for items, but rather that position information can be derived from the associative links (it must be remembered that the model was designed primarily to account for serial recall of short lists).

In what they consider to be a critical comparison of associative and serial-position coding models, they manipulated the ISI for items (ISIs of .5 sec. or 6.5 sec.). They argued that a position coding model would predict better order coding in the "spaced" condition than in the "massed", and that associative models predict the opposite, since closeness in time should facilitate item-to-item associative-processing (see Glenberg model above for similar arguments).
Performance in a serial-recall task was more accurate in the massed conditions in three experiments. While such a finding is inconsistent with a true time tag model, such as Treisman's biological-clock model, it may not be inconsistent with models such as Johnson's (1970) or Estes (Lee & Estes, 1981). In order to form chunks in these models, items would presumably have to co-occupy working memory, and 6.5 sec. ISIs may, in fact, tax the ability to retain items for chunking and position-coding processes.

An interesting feature of the model is the inclusion of beginning and end of list nodes (or markers) to account for performance. These markers might serve critical functions in temporal judgments by serving as anchors in the judgment process (see Glenberg & Fernandez, 1988, also see below).

Comparison of Chunking and Chaining Models. While both chaining and chunking processes provide mechanisms for the preservation of temporal order information, by insuring that an item is integrated into the representation in terms of its input position (localized rehearsal patterns), there are several reasons to believe that chunking strategies would be the more efficacious of the two if temporal order judgments were required. In order to facilitate understanding the arguments which follow schematic representations of simple chaining and chunking models are presented in Figure 1.
CHAINING MODEL

CHUNKING MODEL

LEVEL-2
CHUNK CODES

LEVEL-1
CONTROL ELEMENTS

FIGURE: 1
COMPARISON OF CHAINING AND CHUNKING MODELS
In the chaining model an item's position in the series is directly preserved by the associative links with the preceding and following items, and in no other way (Shiffrin & Cook, 1978; Lewandowsky & Murdock, 1989). The chunking model proposed by Estes (Lee & Estes, 1981) is represented in the schematic diagram shown above. In this model, while direct links among items are not assumed, they are indirectly linked through associations between the control elements at level one. Further, they are also linked indirectly with other items in the chunk through the chunk code itself, and the relative position of this chunk in the series is contained at level 2 through the links between chunk codes. These chunks can also be conceived of as discrete episodes within a longer series. (It should be noted that certain encoding strategies might also contain elements of both chaining (direct item to item links), and chunking (higher-order segmentation)).

In the serial-position judgment task an item's position relative to the beginning of the list must be determined. In the chaining model the subject must engage in implicit serial recall from the beginning of the list and count the number of links until the target item is recovered, and would presumably convert the number of links recovered into an estimate of position.

In the chunking model, on the plausible assumption that subjects create chunks of equal sizes, the judgment would be made by recovering the target item's chunk code and determining the position of this chunk relative to the first chunk in the series. If within-chunk
position is also recovered then an accurate judgment of the target's position can be generated. Further, even if within-chunk position is forgotten, a reasonably accurate position judgment can still be made (i.e., the maximum deviation from true position would be one in this scheme). The relative list length at Level 2 is shorter than at Level 1, and therefore the number of codes that must be preserved in an ordinal fashion in order to allow a reasonably accurate position judgment is smaller (e.g., three as opposed to nine in the chaining model).

There are other reasons why a chunking process might provide a better representation of position information than a simple chaining process. In chaining and chunking models such as the ones shown above, an item's position in the series is contained directly in the associative codes formed between the item and its temporal neighbors. In the case of the chaining account, position is preserved in only two codes, namely the associations formed with the immediately preceding and following items. In the chunking model, with chunk sizes of three items, the target is also linked, through the control elements and the chunk code, to only two items, and so they are equivalent in this regard. However, chunk size should be optional to some extent (Lee & Estes, 1981), and so if chunks are formed containing more than three items (for example Lee & Estes, 1981, induced chunks of four items) than this would provide additional links with temporal neighbors. Therefore, depending on chunk size the number of links with temporal neighbors might increase relative to the chaining process.
It should be noted that larger chunks might increase the likelihood that within chunk position will be forgotten, and therefore might adversely affect performance if serial recall is required. However, if the task required a serial-position judgment for a single item, then by virtue of the fact that the item is being linked with more temporal neighbors, and that the number of chunks is fewer, a reasonable estimate of position might actually be facilitated by creating larger chunks.

Next consider the case where a paired-recency judgment is required for two items. In the chaining model, either a backward or forward implicit serial retrieval is required, and the subject must once again traverse the links until either the later item, or perhaps both items, are recovered. In the chunking model this process might occur at Level 2. Once again, the hierarchical organization in the chunking model should make the task somewhat easier. In that case, if the targets came from different chunks then a decision could be made at this level of organization.

Interestingly, the standard findings in this task would seem to indictate that subjects actually do make use of higher-order segmentations in making paired-recency judgments. In this task a ubiquitous finding is that when the position of the more recent item is held constant, as the lag of the less recent item from the more recent one is increased judgments are both more accurate, and are made more quickly (Fuhrman & Wyer, 1988; Yntema & Trask, 1963; Estes, 1986). This suggests that both items are accessed when making the
recency judgment (since the lag between the two affects performance above and beyond the lag of the more recent item from the end of the list), and this finding poses a serious problem for the chaining model. Within the chaining model, on the assumption of implicit retrieval, increasing the lag between targets should increase the time it takes to make the judgment, since the further apart the items are the more links that must be traversed in order to recover both items, and of course, reaction time data indicates that just the opposite is true. While, it is not necessarily the case that accuracy should decrease with increased lag in the chaining account, there is no clear reason why it should increase.

The distance effect can be more plausibly accounted for in the chunking model. On this account the increase in lag makes it less likely that the target items will contain the same chunk code, and this would allow a decision to be made at the chunk level instead of requiring the recovery of within chunk information (Estes, 1986). Assuming that within-chunk judgments require an extra step (if the chunk codes are the same then the subject must use within-chunk order codes to make the judgment, and these judgments should take longer than with between-chunk judgments). Further, if it is assumed that within-chunk position is more quickly lost than position information at the chunk level (this is predicted from the Estes (1986) model) than the accuracy advantage with increased lag can be accounted for as well.
In short, increased lag benefits paired-recency performance by increasing the likelihood that targets will come from different chunks. However, while this account predicts a big improvement beyond distances of two or three, and this is obtained (Yntema & Trask, 1963), it does not easily account for the fact that the benefit of increasing lag continues to accrue even at relatively longer lags. It is probably the case the supplementary assumptions would be needed to account for this effect. Nonetheless, of the two models the chunking model is more congruent with the distance effect. Thus, along with the argument that chunking models would provide better temporal structure, it is also the case that these models can provide a better account of empirical findings, suggesting that subjects are, in fact, creating some kind of higher-level segmentation as the list is being studied.

To summarize the key points from this section, both the chaining and chunking models assume that localized, as opposed to displaced rehearsal patterns, are crucial in temporal order learning. An item's position in a series is preserved through links with temporal neighbors, and possibly through higher-order structures as well. Given the lack of support for true time tag accounts, and strength theories, and given that in most list learning paradigms extra-list contextual information would be minimal, these kinds of relational coding schemes would seem to be the best candidates for a model of temporal-order learning.
An unresolved issue concerns whether additional processing such as displaced rehearsals for associatively related items from different portions of the list would facilitate (as predicted by Tzeng and Cotton 1980), or interfere with order learning (as predicted in the chaining and chunking accounts). In the following section a model will be proposed that incorporates the idea that chunking facilitates order learning, and from this a consideration of the kinds of encoding strategies that should facilitate temporal-order learning will be given.
CHAPTER III
A MODEL OF TEMPORAL CODING AND
TEMPORAL JUDGMENT PERFORMANCE

Based on the preceding review of the pertinent literature a set of assumptions concerning the encoding and decision processes involved in temporal-order learning can be generated. The first three assumptions refer to the encoding processes that result in a representation of temporal order. To anticipate, it is proposed that encoding strategies that induce localized patterns of rehearsal (linking items that are temporally contiguous) are centrally important in temporal coding. Further, higher-level grouping should be particularly efficacious to this type of learning. These notions are predicted on the assumption that there is no true time tag.

Finally, this account of temporal-order learning predicts that different types of cognitive strategies will be differentially effective in this regard, and therefore it is inconsistent with the idea that temporal coding is an automatic process (Hasher & Zacks, 1979). The second set of assumptions concern the decision processes used when making temporal judgments.
Assumption 1. There is no invariant temporal attribute, or time tag.

This assumption is critical in establishing the ideas which follow. If, in fact, time of occurrence was encoded in some direct, perhaps interval level fashion, as for example the current readout from some sort of biological clock (Treisman, 1963; Bindra, 1976), or as position in a spatial image (Murdock, 1974), then understanding performance in a temporal judgment task would be a most simple matter. Such information would mediate any task where temporal codes might be useful (temporal-order judgments, duration judgments, recall, or recognition). If relatively rapid forgetting of such information is assumed, that could account for lag effects in order judgments, and the recency principle in recall. However, a review of the literature strongly argues against the notion of an invariant time tag.

As mentioned above, various extra-list contextual associations seem to mediate temporal judgments, as well as associations among input items (Guenther & Linton, 1975; Tzeng & Cotton, 1980), and there is no reason to assume that they would if an invariant time-tag was available to subjects. Further, every time-tagging model predicts that increasing the ISI between stimulus items should improve temporal-judgment performance, and this is almost always not borne out in data (Greene & Crowder, 1988; Guenther & Linton, 1975; Shiffrin & Cook, 1978). Further, if performance in a paired-recency task is above chance, and is mediated by an invariant time tag, then performance on lag-judgments should be above chance as well, and as noted above performance on lag judgments for unrelated items is
usually at chance levels (Hintzman, Block, & Summers, 1973; Underwood, 1977; but see Proctor & Ambler, 1976).

Those quantitative models like Glenberg's (Glenberg & Swanson, 1986) or Flexser & Bower's (1974), which employ the notion of a time tag, and provide reasonable accounts for most of the standard empirical findings, do not commit to what actually constitutes this information, and allow the possibility that it might be something like contextual information that is correlated with time. Allowing that contextual information might serve this function really argues against the notion of an invariant attribute. (If contextual structure is different from task to task, or condition to condition, and if temporal information is represented in that way, then this implies that predictions about performance must be task, or condition, specific).

Assumption 2: In the absence of temporally relevant extra-list structure, localized patterns of rehearsal, either chaining or chunking, is the necessary conditions for representing the temporal order of events.

Under conditions where a series of stimulus items are presented over a relatively short span of time, with no highly structured extra-list context, the primary type of information that would preserve a given item's position would be those associative links formed with other items from the series. In order to judge the item's position the ordered representation of the entire series would need to be implicitly retrieved, and from this the temporal order could be
reconstructed. The item's position could only be recovered if the links with its temporal neighbors are sufficiently well encoded to embed the item into a particular location in the ordered representation.

On the assumption that such encoding requires attentional effort (i.e., that it is not an automatic or mandatory process), then it follows that the kinds of encoding strategies that the subject employs will directly influence the degree to which a temporally ordered representation is created. In this section, some of the strategies that have been identified in the memory literature (McDaniel & Kearney, 1984; Shaughnessy, 1981; Weinstein, Underwood, Wicker, & Cubberly, 1979) will be discussed in terms of the degree to which they might preserve temporal-order information.

As noted above, serial-chaining or chunking processes would be the key to linking a given item with its temporal neighbors, and therefore it can be stated that any strategy that includes these elements would be superior. Further, of the two, chunking strategies, by providing a second level of structure, which reduces the number of elements to be ordinally linked at that level, should be superior. Also, if chunks larger than three items are formed, then there are two reasons why this might facilitate temporal judgments. First, this reduces the total number of chunks or groups in the series, and second, it provides a mechanism whereby an item is linked (through the chunk code) with more than two of its temporal neighbors (with chaining direct links are formed with only the preceding and following items in
Obviously, simple chaining or chunking strategies themselves would prove sufficient conditions for temporal-order learning, with chunking the more efficacious of the two. Underwood and Zimmerman (1973) compared performance on two types of lists in a serial-anticipation task. The sixteen-item lists contained four words from four distinctive associative categories. Performance after one study trial was more accurate when the items were blocked according to category, compared to a condition where they were presented in random order. The blocked condition should induce localized as opposed to displaced rehearsal strategies, as well as higher-order segmentation or chunking, and apparently one or both of these factors facilitated order learning. As noted above, Tzeng and Cotton (1980) also found that blocking a categorized list resulted in generally better performance in a paired-recency judgment task.

More elaborative strategies that would induce chunking or grouping, and localized rehearsal patterns, should also facilitate this type of learning. For example, interactive imagery strategies, wherein items are converted into images, and several of these images are combined into a single complex image, would create distinctive chunks or groups from temporal neighbors (this induces localized as opposed to displaced rehearsal patterns), and like simple chunking, might preserve input order at at least two levels (Bower, 1972; Delin, 1969). An imaginal strategy of this type would contain elements of both chaining (direct item to item links) and chunking (higher-order
segmentation). This type of strategy is, of course, contingent on the use of stimulus items for which images can be formed, such as pictures or concrete nouns (Bower, 1972).

One piece of evidence that would appear to be consistent with the claim that complex imagery can facilitate order learning, is that subjects induced to create complex images from pairs of adjacent items perform better on serial recall tasks than subjects given standard instructions (simply to remember the input order) (Delin, 1969; Harcum, 1975). Of course, to account for this it must be assumed that these complex images are ordered at a higher level (of course, pairing would reduce the number of items to be ordered at this level in half). The process of creating complex images from pairs, or groups, of adjacent items induces localized rehearsal patterns, which, it is being claimed here, is a necessary condition for order learning. It would also follow from the above mentioned arguments, that the more items that are integrated into a single image, the shorter would be the list at the group or chunk level. If the subject could remember in which complex image a target item occurred, and if the order of these images has been preserved, then a reasonable estimate of the item's position could be derived. In short, forming complex images (especially ones containing more than two items) satisfies both the localized rehearsal and chunking conditions, and consequently should facilitate temporal-order learning.

It has also been documented that some subjects, when presented with a random word list, will attempt to make up a story (or stories)
that contains the items (Weinstein, et al., 1979; Bellezza, Richards, & Geiselman, 1976; McDaniel, & Kearney, 1984; Shaughnessy, 1981). Such strategies should also induce both localized rehearsal patterns, and chunking processes. Like imaginal strategies they would contain both chaining and chunking components.

The scenes, or episodes, in the story form would contain several of the items interacting in some fashion, and so would qualify as a chunking, as opposed to a simple chaining process (i.e., the attempt to make a coherent story from the items would induce subjects to directly link several, as opposed to two, items into a given episode). The only ambiguity is whether items would be integrated into the story in such a way that input order is preserved. There are two reasons to believe that is the case. First, the order of items in the list constrains the story form; that is the specific story created is determined by the temporal input order. Each new item would add new elements to the story that would change the scenes or episodes in some way, and thus the item's position in the story schema should reflect its position in the input order.

Second, in several studies where subjects were actually instructed to construct stories in order to remember the items, subsequent serial recall performance was superior to other conditions (Bower & Clark, 1969; Weinstein, et al. 1979). Bower and Clark (1969) instructed one group of subjects to create stories from the words in the list, and a second group simply to learn the list. Results from a delayed serial recall task indicated that the former group performed better
than the latter group. Weinstein et al. (1979) instructed one group of subjects to use the method of loci mnemonic strategy (placing each item from the list in a different imaginary location). The second group of subjects also used method of loci, but were also instructed to construct a story to link the items stored at these different locations. This group did better on a subsequent serial recall task than did the first group. Both these studies converge to suggest that story strategies do, in fact, preserve the input order of the items. In that they would also induce elaborative chunking processes and localized rehearsals (items are likely to be integrated into the current scene or episode, as new elements) they would be expected to facilitate temporal judgment performance as well.

Black and Bower (1979) had subjects read short stories and found that different episodes within stories appear to be stored in memory as discrete chunks. The probability of recalling a proposition from a given episode was uninfluenced by the number of propositions in other episodes, but was influenced by the number of propositions within that episode. Interestingly, the more propositions contained in an episode, the greater the probability of recalling any one. Apparently, the larger the chunk sizes can facilitate retrieval processes. There is no reason to assume that subjects, when creating stories from random word strings, would not similarly create discrete scenes or episodes.

To further illustrate the notion that a story strategy would induce order preserving chunking as opposed to chaining processes,
consider the following example. The subject is shown the following words in this order: cat, pie, tree, horse, mouse. The subject might create the following story: A cat ate a pie while sitting in a tree (the subject might even form an image of the scene). A horse with a mouse on his back attacked the cat. The items cat, pie and tree, are integrated into a coherent scene (that is they are directly linked or chunked). The horse and the mouse are similarly linked directly, and their temporal relationship with the first three items is preserved by the fact that they came later in the story. Thus, making a coherent story would force subjects to directly or indirectly relate several items, and the order of events would reflect input order. If the subject attempted to remember the input position of any item, they could access the story form to get a good approximation.

On this localized rehearsal and chunking account there are also several types of strategies that should inhibit order learning as well. Obviously, displaced rehearsal for related items, when these relationships are uncorrelated with input order would induce chunking or clustering by category. However, these chunks would contain no information about input order, in that the pattern of rehearsal is very non-localized. Therefore, displaced rehearsal patterns likely in clustering strategies should inhibit the process of representing temporal order.

On the surface, counting strategies (assigning a number to each item that represents its position in the list) would appear to be a good temporal-order learning process (Saufley, 1976). However, in
this case the only link that is temporally relevant is the association between the item and the number. If this is forgotten then there is no other source of information regarding position in the series. The chaining process on the other hand assures two direct links (preceding and following items), and a chunking process would include links with at least two, and possibly more, temporal neighbors, through the chunk code, as well as higher-level ordering. Thus, these two processes by creating more than one temporally relevant link, should prove more resistant to forgetting effects than counting processes.

Simple maintenance, or rote, rehearsal strategies, where relational coding among input items is minimized, should result in very little information regarding an item's input position getting encoded. It should also be noted that any other item-specific processing (processing that does not induce encoding an item's relations with other items) should inhibit temporal order learning. For example, if the subject is asked to rate the stimulus item in terms of its pleasantness, then this exclusive attention to the current item would interfere with relational coding (linking the item with its neighbors or associates in the list), and while it might produce better free recall than rote rehearsal, because it induces semantic processing for the item (Craik & Lockhart, 1972), it should prove no more efficacious than rote rehearsal in preserving temporal-order information.

Finally, it should be noted that subjects often report using more than one type of strategy within the same serial list. It might be the case that these shifts in cognitive processes across the list
would qualify as changes in process context (see above), and would result in distinctive higher-order episodes being created. For example, if the first ten items are chunked, and if the subject uses a story for the second ten items, and finally rote rehearsal for the last ten items, then recovering information about the way in which an item was encoded, might allow a reasonable estimate as to where the item occurred in the list (Underwood, 1977; Block, 1982). Still, it would not be expected that performance in this condition would equal conditions where highly localized chunking strategies are used consistently across the list.

Only one study to date has directly manipulated subjects encoding strategies in a temporal judgment task (such manipulations have been employed in serial-recall tasks as noted above). Tzeng (1976) presented subjects with a categorized word list (10 categories with five exemplars of each presented in random order). One group of subjects was told to rehearse associatively related items aloud, and a second group was told to engage in maintenance rehearsal (simply repeat aloud the item currently on the screen). Subjects in the relational condition were therefore induced to engage in a good deal of displaced rehearsal while the other subjects did item-specific rehearsal, and performance on a surprise serial-position judgment task did not differ significantly as a function of strategy. Tzeng, a proponent of displaced rehearsal theory (see above), argued that the failure to obtain performance differences as a function of strategy indicated that the necessary relational coding processes (he assumes
these to be both localized and displaced rehearsal patterns) are automatic or mandatory.

However, given the above mentioned arguments that localized rehearsal patterns are necessary conditions for temporal-order learning, it could also be argued that both displaced rehearsal patterns and maintenance or item-specific processing inhibit, to a similar extent, these localized rehearsal patterns, and that this is why no performance differences were obtained. If more appropriate strategies had been included in the experiment, then performance differences might have been obtained. Experiment 3 in the current study addresses this issue (see below).

The foregoing discussion of appropriate temporal-order learning strategies has really been predicated on two key assumptions. First, that there is no invariant time tag, and second, that the kinds of localized rehearsal patterns necessary for this type of learning are not mandatory or automatic processes (Hasher & Zacks, 1979; Tzeng & Cotton, 1980). The data suggesting strategic influences on order learning have been obtained in tasks where serial recall is required. However, it is not necessarily the case that serial-position or paired recency judgments are mediated by the same information. Therefore, it must be determined whether strategy effects in line with the current model, are obtained with these latter dependent measures.

Retrieval and decision assumptions. In this section some assumptions regarding the processes involved in making temporal judgments will be described.
Assumption 3. Temporal judgments about single target items (serial position or absolute judgment of recency) or pairs of target items (paired recency or lag judgments) are made by trying to determine the item's temporal location relative to a reference point, usually the beginning or the end of the list.

Evidence from subject protocols, serial position curves, and differences between tasks in terms of the degree to which they produce primacy or recency effects, converge to support the notion that when an item is presented in a judgment task subjects attempt to ascertain its position relative to a salient temporal reference point. This notion contrasts with the assumption, found in time tag models (Treisman, 1963; Bindra, 1976) that absolute temporal information is represented and used to make temporal judgments. On that account, position judgments would be made by simply retrieving the temporal attribute. As noted above, on the associative coding account the position of a target must be determined relative to the beginning or the end of the list, by engaging in implicit retrieval of the items, or chunk codes, that fall between the target and the endpoint (recall that in a chaining account the items would need to be counted, and in a chunking account the process might occur at the chunk level itself). From this type of process the distance of the item from the endpoint can be determined, and a number can be assigned.

Underwood (1977) asked subjects to report the strategies that they used when making serial-position judgments, and most reported trying to determine the item's position relative to the ends of the list.
Thus, subject's reports are consistent with this notion. In the experiments reported in the current study, it was observed that subjects, if free to make serial-position judgments for items in any order they chose, tended to respond to early and late items and work their way in to the middle of the list. This again might suggest that the ends of the list serve as reference points for temporal judgments.

As noted earlier serial-position judgments usually show strong primacy effects (Hintzman, Block & Summers, 1971). When absolute judgments of recency are made, wherein a subject must determine how far back in the list a target item occurred, recency effects are found (greater accuracy for items near the end of the list), but primacy effects often are not obtained (Wells, 1974; Tzeng 1976). Both tasks require temporal judgments about single items, but serial-position judgments may encourage using the beginning of the list as a reference point (since subjects are assigning numbers, with early positions receiving lower numbers). Absolute recency tasks, on the other hand, might encourage backward scanning; with their using the end of list as a reference point, since subjects must assign a number that represents how far back in the list the target occurred (see Hacker, 1979, and Muter, 1980, for reaction time results suggesting that absolute-recency judgments involve backward search in STM tasks).

In the serial-position task, using the beginning of the list as a reference point would tend to keep items from the beginning of the list more active, and therefore less likely to be forgotten, while recency judgments would do the same for later items. That could
account for the observed primacy effects in the former task, and recency effects in the latter. Again, this suggests that subjects are using the end of the list as a reference point in making these judgments.

The implication of this idea is that temporal judgments are made with reference to other items within the list, and with reference to salient markers. They probably do not involve the recovery of information concerning absolute position in time.

Assumption 4. Temporal-order information is forgotten fairly quickly under many conditions, and when this direct information is unavailable, subjects may rely on memory trace-strength as an alternative cue. This strength information also may operate as bias, even when temporal cues are available.

If the previously mentioned encoding and decision assumptions are correct, they might allow us to give a reasonable account of performance under different experimental conditions (i.e., when subjects are able to engage in order-preserving relational coding processes, for example). The information that enters the decision mechanism is clearly identified. However, given the fact that forgetting of these codes is often fairly rapid (Underwood, 1977; Hintzman, Block & Summers, 1973; Estes, 1985), it must be determined whether under these conditions, subjects will rely on guessing or will instead use memory-trace strength as a cue. Further, if they do use strength as a cue under these conditions, it should be determined whether this cue also might influence judgments under conditions where
some direct information is available as well, operating as a kind of judgment bias. Obviously, before this latter question can be addressed, it will be necessary to establish that subjects are sensitive to memory-trace strength, and do, in fact, use it as a temporal cue, in certain conditions.

The data to this point are equivocal. In autobiographical memory research Brown, Rips & Shevell (1985) found that when paired-recency judgments were made about well known events, those events for which subjects could be presumed to have more detailed memories, tended to be judged as more recent than those that were somewhat less well known. They interpreted this as reflecting a kind of strength bias in recency judgments.

A number of researchers have found that when one of the targets in a paired-recency judgment is presented twice (repeated) and the other one presented once, the repeated item is often judged as more recent (Peterson, Johnson & Coatney, 1969; Brelsford & Freund 1967). However, as noted above, this effect has been questioned, because of the lack of support for the notion of single-trace summation, as well as the fact that the reverse effect is obtained when primacy judgments are made, and that is a result not predicted from the strength account (Tzeng & Cotton, 1980; Flexser & Bower, 1974).

Block (1982) found that when subjects were given a levels of processing orienting task (shallow or deep), and were subsequently required to judge in which of two study lists a given item occurred, there was a tendency to place items that had undergone shallow
processing into the earlier list, and items that had received deep processing into the later list. That might suggest a strength bias operating in the task. That follows from the idea that items encoded less deeply would presumably have "weaker" memory-trace strength.

However, Wells (1974), found that subjects confidence ratings in a recognition task were not predictive of performance in an absolute judgment of recency task. Presumably, on the strength account, items with higher confidence ratings would have stronger memory-trace strength, and should have been judged to be more recent. However, Wells used fairly short lists and under these condition subjects probably would have more direct information available to the decision processes.

Wolff (1966) used CVC's (consonant-vowel-consonant items), that were either high or low in associative value, in a mixed list design. In the test phase subjects were required to make paired recency judgments for four types of test pairs: high-high, high-low, low-low, low-high. Wolff predicted, on the strength account, that if strength were greater for high associative value CVCs, then performance should be particularly good in the low-high pairs and particularly poor in the high-low pairs. He did find that performance was well below chance in the high-low condition, but did not find an advantage for the low-high pairs so the data provide only limited support for the strength bias notion. However, the lists were fairly short, and presentation rate was slow (4 seconds), which means that some direct cues may have been available to subjects, and they would diminish
strength effects. This study is quite similar to one of the experiments reported below, and will be discussed later.

In summary, while it seems plausible to posit a role for strength of the memory-trace as a possible secondary cue to recency, there is only limited support for this idea in the literature. Since this type of cue is still proposed in recent models of memory (Gillund & Shiffrin, 1984; Hintzman, 1988) its possible role in temporal judgments needs to be more carefully investigated.

Summary of encoding and decision assumptions. It was proposed that under conditions where extra-list context is minimal, localized rehearsal patterns, chaining or chunking, are necessary conditions for representing an item's position in the series. No direct interval-level temporal information is assumed. The encoding processes are thought to require attentional effort, and consequently different types of encoding strategies should be differentially effective in temporal-order learning. Judgments are made by trying to locate the position of an item(s) relative to the beginning or end of the list, by implicit retrieval of intervening items or chunk codes. Under conditions where this information has been forgotten, it may be that subjects will convert an estimate of memory-trace strength into a judgment of temporal position.
A number of researchers have claimed that the encoding of certain types of information may be automatic, such as frequency of occurrence, time of occurrence (temporal coding), spatial location, and access of word meaning (Hasher and Zacks, 1979, 1984). Hasher and Zacks proposed six criteria that an encoding process must meet if it is to be classified as automatic. They argued that the process should be unaffected by intentionality, cognitive strategy, practice effects, the simultaneous operation of other attentionally demanding non-automatic processes, organismic states that alter arousal level, and developmental factors (age). It is also assumed by Hasher and Zacks that automatic processing is optimal; that is, it would not be improved by intentional factors, but that idea has been seriously challenged recently on logical grounds (Logan, 1988; Sanders, Gonzalez, Murphy, Liddle & Vitina, 1987).

Controlled or strategically-driven processing can be defined operationally, as those processes that do not satisfy these automaticity criteria (Hasher & Zacks, 1979). While these two types of processing have been generally viewed as discretely different, there has been a growing tendency for researchers to argue that there may be a continuum of processing, with some processes falling closer to automaticity and others requiring more attentionally-controlled effort (Logan, 1988; Naveh-Benjamin, 1988; Sanders, et al. 1987). This shift from the dichotomous view has come about primarily as a result of a growing body of evidence which suggests that each of the
above mentioned encoding processes fail to strictly meet the six criteria outlined by Hasher and Zacks (1979) (Greene, 1986; Sanders et al., 1987; Hanson & Hirst, 1988; Naveh-Benjamin, 1988).

The implication of the idea that certain memorial processes operate automatically, is that they have special status in the information processing system. They can be thought of as self-contained modules (Fodor, 1983), whose operations are uninfluenced either by other cognitive operations, or by higher-level intention. Based on this idea, a certain kind of working memory architecture is implied; one with many parallel processing modules operating to encode these different types of information. For example, several recent models have incorporated specialized temporal coding mechanisms that automatically encode the temporal context in which an item is embedded (Gillund & Shiffrin, 1984; Detweiler & Schneider, 1988). These approaches are predicated on the hypothesis that temporal coding is uninfluenced by intentionality or idiosyncratic strategies. It therefore becomes quite important to examine closely the claim that temporal coding is an automatic process.

**Evidence for automaticity**

The primary piece of evidence that has led to the speculation that temporal coding is an automatic process is the failure to find any difference in performance between subjects instructed to encode order (intentional) vs. subjects given no such instructions (incidental). That is, subjects do no better when they are expecting an order test
than when they are not. Toglia and Kimble (1975) instructed one group of subjects that a serial-position judgment task would follow the study list, and told a second group simply to expect a memory test (the test was not specified). They found no difference in performance as a result of instructions. This finding was replicated by Tzeng (Tzeng, 1976), and Pugh (1986).

Other evidence suggesting automaticity has been obtained as well. Underwood and Zimmerman (1966) found that presentation rate did not affect performance on a serial-position judgment task, whereas study time has well documented effects on other types of memory tasks such as recall. McCormack (1981) found no evidence of age differences in temporal judgment performance, whereas age differences do emerge in a variety of other memory tasks (Hasher & Zacks, 1988). As noted earlier Tzeng (1976), found that one group of subjects induced to engage in relational coding did no better on serial-position judgments than a group forced to engage in a restricted maintenance rehearsal strategy. (However, if the assumptions of the current model are correct then that result might be explained by the notion that neither strategy allowed for localized rehearsal patterns, and consequently were equally poor).

Evidence against automaticity

There also have been a number of studies reported recently which have provided evidence that appears inconsistent with automaticity accounts. As noted above, Pugh (1986) reported data which suggests
that attending to semantic or associative relationships that are orthogonal to input order (displaced rehearsals) interferes with temporal-order learning as measured in a paired-recency task, and this would not be expected from the automaticity hypothesis.

Other evidence more consistent with the controlled-process account has been obtained as well. Type of materials (e.g., concrete nouns vs. abstract nouns or nonsense items) affects accuracy of both paired-recency and serial-position judgments (Yntema & Trask, 1963; Michon & Jackson, 1984). Subjects are more accurate on concrete word lists than on abstract lists (even when forgetting is controlled). Since concrete words might better lend themselves to relational coding strategies such as interactive imagery and stories, the advantage might be interpreted as suggesting strategic influences (Michon & Jackson, 1984).

Michon and Jackson (1984) presented a random series of pictures containing items from two different associative categories. In one condition, items from one category appeared at the top of the screen, and items from the second at the bottom, although presentation order was randomized. In a second condition the same order was used but all items appeared on the same position on the screen. The authors assumed that categorical clustering would be more likely in the former condition, and that such clustering might interfere with learning the order of the series as a whole. Consistent with that prediction, they found that within-category paired-recency judgments were more accurate, and between-category judgments less accurate, in the
separated than in the uniform condition. This finding is similar to the categorical clustering effects reported by Pugh (1986), and suggests that attending to associative information, when this information is uncorrelated with input order, inhibits order learning.

Adding a background extra-list context facilitates the encoding of order information for target items from a serial list. For example, Guenther and Linton (1976) (see above) reported that subjects did better on paired-recency judgments for a list of unrelated pictures when a chronologically ordered story was heard simultaneously, even though the story bore no relation to the target pictures. This extra-list context appears to facilitate the encoding of order information, and it is not clear that an automaticity account would predict this.

Finally, while at odds with all other studies in the literature that manipulated instructions, Michon and Jackson (1984) did obtain an advantage on a surprise position judgment task for subjects expecting free-recall over those expecting a recognition task. They argued that the encoding strategies in the two conditions differed, and that this directly influenced temporal coding performance.

However, it could also be argued that subjects expecting a recognition test might not sufficiently attend to the items to activate the automatic coding processes, and advocates of the automaticity hypothesis do assume that a certain amount of initial attentional effort is necessary to initiate these processes, which then operate automatically (Hasher & Zacks, 1979; Logan, 1988).
Therefore, to unambiguously refute the automacity hypothesis it would have been necessary to assure that subjects in both the expect recognition and expect recall conditions were employing sufficient intitial encoding effort, and this was not done.

**Intentionality Effects**

There is a growing body of evidence that casts doubt on the claim that temporal coding is a truly automatic process. However, a truly controlled process should be to some extent optional, and the fact that subjects expecting an order test do no better on the test than subjects not pre-informed is therefore somewhat troubling for this account. It could well be, that there is a continuum of processing with truly automatic and truly controlled processes falling at the two ends of the continuum (Logan, 1988; Sanders et al., 1987; Naveh-Benjamin, 1987). Effects such as interference from categorical clustering, or differences in performance as a function of expectation of a recognition versus a recall task, would suggest that temporal coding is not a strictly automatic process. However, the failure to obtain intentionality effects when instructions are manipulated, might suggest that intention to encode order information has no influence on the processing, and this would imply that temporal coding falls closer to the automaticity end of the processing continuum than to the controlled process end. Further, in that it might be expected that subjects in the intentional condition might very well employ somewhat different encoding strategies than subjects in the incidental condition, that would suggest that encoding
strategies have no influence as well. That, in turn, would be inconsistent with the ideas outlined in the previous section.

If, on the other hand, encoding strategies do affect temporal-order learning, as suggested in the current model, then the failure to obtain expectancy effects must imply that subjects in the two conditions are employing either the same, or equivalent, strategies in both cases. There are a number of reasons why this might be so. One possibility is that subjects in the intentional condition simply do not know which kinds of strategies are more efficacious in order learning. This implies a lack of knowledge about how memory operates (knowledge of how memory operates is called metamemory, and there is a large and interesting literature exploring the issue of how well subjects understand their own memorial processes (see McDaniel and Kearney, 1984, or Shaughnesssey, 1981 for reviews of this literature).

If metamemory is poor, then it may be that simply informing subjects that a temporal judgment task will be given will not induce them to employ better strategies. However, if instead of merely being informed, subjects were given training with the task itself, the experience might improve their understanding of what kinds of encoding strategies are most effective. In that case if their performance was compared with subjects given no training with the task, strategic differences might emerge. It should be noted, in this regard, that in both spatial memory tasks, and frequency of occurrence tasks, instructional effects also are not obtained (which led to the speculation that these types of encoding processes are automatic).
However, in each of these literatures, other, more direct, manipulations have revealed strategic influences (Naveh-Benjamin, 1987; Sanders, et al., 1987; Greene, 1985). Again, this would suggest that merely informing subjects about the nature of the test is not a strong enough manipulation to reveal strategic influences.

It might also be the case that subjects do possess knowledge of the kinds of strategies that result in good temporal-order learning and that subjects choose these strategies when pre-informed about the test. However, subjects in the incidental condition may simply choose the same, or equivalent strategies because they believe these will facilitate memory performance in general. In other words, the failure to obtain intentionality effects may arise because subjects in both groups tend to choose more efficacious strategies independent of instructions.

In order to distinguish between these different possible accounts of the failure to obtain intentionality effects, it would be useful to create conditions where expectancy is manipulated, and then determine what kinds of strategies subjects actually do employ in these conditions. Experiments 1 and 2 of the current study represent an attempt to do just this.
The failure to obtain differences in performance between subjects expecting a temporal-order test (intentional), and those not expecting such a test (incidental), has been interpreted as supporting the claim that temporal coding is an automatic process (i.e., one that is uninfluenced by strategic processes). However, as noted above, there are reasons to doubt whether this manipulation is appropriate to induce strategic shifts (for example, in the spatial coding and frequency coding literatures instructional manipulations failed to affect performance, but in both cases other manipulations had revealed strategic influences). Since the failure to find intentionality effects has really been a crucial piece of evidence supporting the automaticity claim, it is important to find other variables that might, in principle, induce subjects to vary the extent to which they attend to temporal order, and determine if performance on a temporal-judgment task is significantly affected by these manipulations.

The first two experiments in this series manipulated prior experience, in order to determine whether strategic effects are obtained as a function of whether subjects are, or are not, given
experience with a temporal judgment task on a training trial (one group of subjects is trained with a serial-position judgment task, and a second group is trained with a free-recall task, then on the second critical trial both groups are given a serial-position judgment task). With this manipulation subjects trained with a temporal judgment task not only be should be primed to attend to temporal order on the second trial, but their initial experience might provide the necessary feedback to allow them to adopt more appropriate encoding strategies on that second trial.

Further, an analysis of the kinds of strategies that subjects actually do employ in these conditions was conducted. This would allow two issues to be addressed. First, it might help determine whether different types of strategies are differentially effective in temporal-order learning in general, as suggested in the present model. Second, if differences in performance as a function of chosen strategies are obtained, then the extent to which these chosen strategies vary as a function of the training manipulations can be assessed. In this way the locus of any effects obtained from the training manipulation can be more precisely determined.

The first experiment in this series was designed to manipulate the degree to which subjects attend to the temporal-order of the stimulus items by manipulations of training and type of materials. The second experiment replicates some of these conditions, and also includes an analysis of subjects reported encoding strategies in these conditions. Finally, a third experiment directly manipulates encoding strategy to
determine if performance differences are obtained.

Experiment 1

It was reasoned that informing subjects that a temporal judgment task will follow the study list might not be sufficient to induce subjects to employ more efficacious order learning strategies (especially if subjects knowledge of good strategies is not well developed). However, prior experience with a temporal-judgment task might provide the necessary training or feedback to induce subjects to shift to more appropriate strategies. Further, the type of materials studied might have an influence on which type of encoding strategy is employed (for example, certain materials might induce subjects to engaged in more localized rehearsal patterns, while other materials might encourage displaced rehearsals). If strategies directly influence temporal-order learning then these conditions might reveal these influences more clearly than would a simple intentional/incidental manipulation.

In this experiment all subjects were given two study-test trials. The effect of Trial 1 experience on Trial 2 performance was the primary concern. The specific manipulations used were as follows:
1) The Trial 1 study list was either composed of CVCs (consonant-vowel-consonant nonwords) or three letter nouns.
2) The memory test on Trial 1 was either free-recall or a serial-position judgment task.
3) The Trial 2 list was either composed of CVCs or three letter nouns. The test which followed the Trial 2 list was always a serial-position judgment task.

Predictions. It was reasoned that subjects, given a serial-position judgment task on the first trial, would, on the second trial, be more likely to employ order-preserving strategies, than subjects given a free-recall task on the first trial. To the extent that this would affect the efficacy of temporal coding, the first group of subjects should perform better on the Trial 2 serial-position task than the second group.

The CVC vs. word manipulation on the training trial was included to determine whether the type of materials studied on Trial 1 would influence the type of strategies that subjects employ on Trial 2. A list of words may contain many items that possess pre-experimental associative relationships (for example, items might come from similar categories, such as living thing, food, etc.), and if presentation order is uncorrelated with these relationships then displaced rehearsal patterns would be likely. However, nonsense items would presumably contain items that have no pre-experimental associative relationships. With CVCs, then, more localized rehearsal patterns would be encouraged (such as chaining or chunking). If these strategies are carried over to the second trial list (set effects), then it might be predicted that subjects trained with CVCs would do better on the Trial 2 serial-position judgment task. This follows from the assumption, outlined above, that localized as opposed to
displaced rehearsal patterns are necessary conditions for temporal-order learning.

However, order learning might, in general, be superior for word lists, because more elaborative encoding schemes would be possible for meaningful material (Yntema & Trask, 1963; Michon & Jackson, 1984). For example, subjects might use imagery or stories to facilitate order-coding processes for word lists, and these would be unlikely for CVC lists. Performance on the Trial 2 serial-position judgment task would test this, in that performance on word lists and CVCs can be directly compared.

The pattern of interactions between materials and prior test might reveal the combined influences of these strategy manipulating variables. Specifically, it would be predicted that subjects given CVC lists and position judgments on Trial 1 and word lists on Trial 2 would perform best, while those given word lists and free recall tasks on Trial 1 and CVC lists on Trial 2 would do worst.

The serial-position judgment task was used for the following reasons. First, in that it requires a more precise estimate of a target’s position than the paired-recency task, it might prove more sensitive in revealing differences due to rehearsal patterns (e.g., localized vs. displaced). Second, several previous studies which failed to obtain expectancy effects used precisely this task (Toglia & Kimble, 1976; Tzeng, 1976; Zacks et al., 1984).
Method

Design. A 2x2x2 between-subjects design was used. The variables were Trial 1 test (free recall vs. position judgment); Trial 1 materials (CVC's vs. words); and Trial 2 materials (CVC's vs. words).

Materials. Four lists of 30 items each were constructed: two were word lists containing three letter nouns, and their mean frequencies were equated using the Kucera and Francis (1967) norms. In addition, there were two lists of CVCs taken from Archer (1968) and the mean associative value was equated for the two lists. Three random orders were generated for each list and used equally often across subjects. Slides were created for each of the lists, and they were mounted in Kodak slide carousels according to the appropriate randomizations, and presented to subjects at a rate of 5 seconds per item.

Test lists were composed by randomly listing the thirty items in a single column on a piece of paper with a blank next to each word. Two such lists were composed for each of the four study lists (and with the three presentation orders used for each study list this resulted in six different test list orders for each of the lists).

Subjects. The subjects were 80 undergraduates enrolled in a psychology course at the Ohio State University. They received partial credit toward fulfilling a course option. Subjects were tested together in groups ranging from four to eight members.

Procedure. Subjects in each condition were told that they would be studying a list of items, and after which they would be given a memory test (the specific nature of the test was never specified).
After viewing the 30 item study list subjects were given a two-minute retention interval during which they filled in the date, condition number, and subject number, at the top of the test sheet.

In the free-recall condition a blank piece of paper was provided and subjects were told to remember as many items as they could, writing them down in any order. They were given five minutes to complete this task.

In the serial-position judgment condition subjects were given the test sheets, which, along with the items and blanks, also included a number line from 1-30 at the top. They were informed that the study list contained 30 items and that these had been listed in a column in a random order. For each item they should assign a number from 1-30 that corresponded to the item's position in the study list. They were further told to try to use each number only once (this is why the number line was made available), and that they should fill in all thirty blanks, even if unsure of their answers. The order in which they filled in their answers was not controlled. As in the free-recall condition they were given five minutes to complete the task.

Results

The data from the Trial 2 serial-position judgment task are presented in Table 1. The dependent variable computed for each subject was the correlation between true position and judged position (then transformed using Fisher's r to Z' transformation). Mean Z' scores are shown in Table 1 for each condition.
When performance on Trial 1 vs Trial 2 was compared for those subjects who were given serial-position judgments on both trials, a significant improvement on Trial 2 was observed $F(1,38) = 6.23$, $p < .05$ (Trial 1 mean = .399; Trial 2 mean = .624). This improvement with practice has been observed in other research (Underwood & Malmi, 1978; Zacks et al., 1984), and is not predicted on the automaticity account (practice effects are indicative of strategically controlled processing according to Hasher and Zacks (1979)). However, given that subjects in the free-recall condition actually performed better on Trial 2 than this group (see below), this practice effect probably cannot be attributed to training with a temporal judgment task itself. Instead it might reflect a general improvement from experience with previous study-lists. Therefore, given the current interests, it will not be considered further.
Table 1: Trial 2 mean Z' scores from Experiment 1

### CVC List on Trial 2

<table>
<thead>
<tr>
<th>Trial 1 List Type</th>
<th>CVC</th>
<th>Word</th>
<th>$\bar{X}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Trial 1 Test Type</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPJ</td>
<td>.418</td>
<td>.275</td>
<td>.347</td>
</tr>
<tr>
<td>FR</td>
<td>.411</td>
<td>.529</td>
<td>.470</td>
</tr>
<tr>
<td>$\bar{X}$</td>
<td>.415</td>
<td>.402</td>
<td>.409</td>
</tr>
</tbody>
</table>

### Word List on Trial 2

<table>
<thead>
<tr>
<th>Trial 1 List Type</th>
<th>CVC</th>
<th>Word</th>
<th>$\bar{X}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Trial 1 Test Type</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPJ</td>
<td>.907</td>
<td>.742</td>
<td>.825</td>
</tr>
<tr>
<td>FR</td>
<td>1.071</td>
<td>1.341</td>
<td>1.206</td>
</tr>
<tr>
<td>$\bar{X}$</td>
<td>.989</td>
<td>1.042</td>
<td>1.016</td>
</tr>
</tbody>
</table>
The Trial 2 data showed a main effect of list type, $F(1, 72) = 23.09$, $p < .001$, indicating that subjects were more accurate on word lists (mean = 1.015) than on CVC lists (mean = .408). There was no effect of the type of list studied on the first trial ($F < 1.0$). A main effect for type of test on the first trial indicated, surprisingly, that subjects trained with free recall did somewhat better on the Trial 2 serial-position judgment than subjects trained with the same position judgment task $F(1, 72) = 4.03$, $p < .05$, (means of .838 vs. .585 respectively), which is exactly the opposite of the predicted effect. No other effects or interactions were significant.

Discussion

The fact that subjects did better on temporal judgments for word lists than for CVC lists appears to be similar to findings reported by Yntema and Trask (1963), and Michon and Jackson (1984). In these studies subjects were more accurate in paired-recency and serial-position judgment tasks with concrete words than with abstract words. One possible interpretation of this effect is that the concrete words used in the current experiment afforded subjects an opportunity to use more elaborative strategies (such as imagery or creating stories), and that this produced the advantage over CVC lists.

Alternatively, the current finding of an advantage for words over CVCs could be simply the result of greater forgetting with CVC lists. However, Michon and Jackson (1984) who controlled for forgetting, still found a significant advantage for concrete over abstract words in a serial-position judgment task, and that would argue that other
factors may be operating to produce the advantage. Also, there is reason to believe that recognition will be high in both conditions with these relatively short lists (Zacks, et al., 1984; Michon & Jackson, 1984; Tzeng 1976). In any event, this result is inconsistent with the notion that an invariant time tag is encoded for items, since the lexical status of an item should not influence this process.

An unexpected finding was that free-recall practice produced better performance on subsequent serial-position judgments than practice with the same position task. Apparently, those subjects primed to focus on order information employed strategies that were less efficient to the task than those employed by subjects not so primed. This effect is consistent with the idea that strategic processing (based on expectancy) can affect temporal coding efficacy. However, while consistent with the assumption that strategic factors influence performance, it is, nonetheless, somewhat confusing. It is also possible that the disadvantage of training with the same task might reflect the operation of proactive interference (Wickens, Born, & Allen, 1963) on the second trial for subjects engaged in the same type of judgment process.

One finding from the literature on frequency estimation may provide some insight into the current effect. Greene (1986) instructed one group of subjects to construct sentences containing target items; and another group was to try to remember the frequency of occurrence for each word; and a third group was to engage in maintenance rehearsal. He found that the group that engaged in the
elaborative strategy did better on a frequency estimation task than did the group specifically attending to frequency (there was no difference between the counting and maintenance rehearsal conditions).

A similar phenomenon might be occurring in the current experiment. Subjects exposed to a serial-position judgment task on Trial 1 may be more likely to try to remember position by assigning numbers to each item, or using some similar tagging strategy. In the preceding discussion on the relative efficacy of different encoding strategies it was reasoned that counting strategies would provide only one temporally relevant link (the association with the number), and that if this were forgotten there would be no other information on which to base the judgment. Chaining and simple or elaborative chunking strategies, on the other hand would result in at least two and possibly more temporally relevant associations being formed, and they might be expected to be superior. The subjects expecting free recall may be more likely to engage in elaborative strategies, such as making up stories that contain the words, and to the extent that these strategies preserve order, this might account for the performance advantage.

In summary, the data from the first experiment did not conform to the initial hypotheses. There was no effect of the type of list studied on the first trial. Recall that it had been predicted that training with word lists might encourage subjects to adopt encoding strategies that contained more displaced rehearsal patterns than training with CVC lists and that this might carry over to the second
trial. That was not confirmed in the data.

It was also predicted that training with a serial-position judgment would be superior to training with a free-recall task, when position judgments were required on the second trial. The opposite pattern of results was obtained. The advantage for free-recall training does suggest strategic effects, but also suggests that subjects do not necessarily choose optimal encoding strategies when expecting an order judgment. This apparent metamemory shortcoming (McDaniel & Kearney, 1987) may suggest a reason why intentionality effects are not, in general, found. Subjects simply may not know what type of strategy leads to optimal temporal coding performance. In any event, strategies will be examined more carefully in the following experiments.

Experiment 2

The second experiment attempted to replicate the finding from the first experiment that, on a Trial 2 position judgment, free-recall training is superior to training with the same temporal judgment task, and to determine the locus of this effect.

It was reasoned that perhaps subjects in the serial-position training condition adopted strategies, such as counting, that were actually less efficacious than the kinds of strategies adopted when subjects were not primed to attend to order information. That would suggest that subjects knowledge of the kinds of strategies that produce optimal temporal-order learning is not well developed.
In order to address this question protocols were collected, in which subjects were asked to describe the type of strategy they had employed on the second trial, and whether these were the same strategies that they had employed on the first trial. Such information might allow for a determination of the locus of the free-recall advantage (chosen strategies might be different in the two conditions). It also might allow a determination to be made of the relative efficacy of different kinds of encoding strategies.

Recall, that the current model predicted that different types of commonly reported strategies should be differentially effective in temporal-order learning. Specifically, it was predicted that strategies that induce localized patterns of rehearsals, such as chaining or chunking should be superior to either those that limit relational coding (such as rote rehearsal, and other item-specific coding processes), or those that induce displaced patterns of rehearsal (such as associative clustering). The collection of protocols might provide a test for these assumptions (it should be noted that to date no research has attempted to determine a ranking of the relative efficacy of different kinds of strategies in temporal judgment tasks, and that is probably due to the pervasiveness of the idea that temporal coding is an automatic process).

Also, in the first Experiment response order was not controlled. That is, subjects could fill in the thirty responses in any order they chose, and it is possible that this might, in some unspecified way, produce the advantage for free recall training. To control for this,
half of the subjects in this experiment were given the same test sheets as in the first experiment, and half were shown the test items one at a time on the screen, and they made a position judgment for each item when it was shown. That allowed the experimenter to control the order of position judgments. It also would be empirically interesting to determine whether response order has an influence on performance.

Finally, since the free recall advantage did not interact with list type (CVC vs. word) in the first experiment, the decision was made to use only word lists in the second experiment.

Method

Design. A 2x2 between subjects design was used in this experiment. The variables were Trial 1 test (free recall vs. serial-position judgment) and type of test format (free vs. experimenter controlled).

Materials. The word lists were the same as those used in Experiment 1. Also the same randomizations were used.

The test lists were also the same as in the first experiment (two randomizations with three study-list orders producing six test-list orderings). However, half of the subjects were given the same format as in the first experiment, where the test items were listed in a column on a sheet of paper, and subjects were free to respond in any order they wanted. The other half were shown the test items one at a time on the screen (in the same orders that they appeared on the test sheets), and they were therefore forced to make a position judgment.
for each item as it appeared on the screen. Thus, response order was determined by the experimenter for this group. The subjects saw the item on the screen, wrote down the word in a blank, and put a number from 1-30 next to it. All other conditions in this experiment were the same as in Experiment 1.

**Subject.** The subjects were 205 introductory Psychology students drawn from the same population as the first experiment.

**Procedure.** The procedure was the same as in the first experiment with the exception that after the second test the subjects were instructed to write down any type of strategies that they may have used when studying the second list, and whether this was the same strategy that they had used on the first trial.

**Results and Discussion**

Once again correlation scores were computed based on the relation between true and judged position and converted into Fishers Z' scores. The data are shown in Table 2.

None of the mean differences approached significance in this experiment (F's all less than 1.0). The pattern of means seems to indicate that when free-recall training and uncontrolled response conditions were used (as in the first experiment) there was an advantage over all other conditions (including position judgment training and uncontrolled response order, which was also used in Experiment 1). However, even with this larger sample size, separate analyses performed on just these two cells, which matched conditions Experiment 1, revealed no significant differences. So while the
pattern of data is consistent with Experiment 1 the effect is not significant. In short there was no reliable advantage for free-recall training, and no effect of response order as well.

Table 2: Trial 2 mean Z' Data from Experiment 2

<table>
<thead>
<tr>
<th>Type of Test Format</th>
<th>Uncontrol.</th>
<th>Exp. Control.</th>
<th>( \bar{x} )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Trial 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Test Type</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPJ</td>
<td>.685</td>
<td>.682</td>
<td>.684</td>
</tr>
<tr>
<td>FR</td>
<td>.760</td>
<td>.696</td>
<td>.728</td>
</tr>
<tr>
<td>( \bar{x} )</td>
<td>.723</td>
<td>.689</td>
<td>.706</td>
</tr>
</tbody>
</table>
### Table 3: Protocol Data From Experiment 2

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Mean Z' score</th>
<th>Prop. Ss. in SPJ</th>
<th>Prop. Ss. in FR</th>
<th>Total Prop.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Story</td>
<td>.947</td>
<td>.22</td>
<td>.18</td>
<td>.21</td>
</tr>
<tr>
<td>Visual Imagery</td>
<td>.890</td>
<td>.04</td>
<td>.07</td>
<td>.05</td>
</tr>
<tr>
<td>Chunking</td>
<td>.690</td>
<td>.25</td>
<td>.14</td>
<td>.20</td>
</tr>
<tr>
<td>Rote</td>
<td>.661</td>
<td>.05</td>
<td>.01</td>
<td>.03</td>
</tr>
<tr>
<td>Mixed</td>
<td>.638</td>
<td>.16</td>
<td>.19</td>
<td>.17</td>
</tr>
<tr>
<td>Not Categor.</td>
<td>.548</td>
<td>.15</td>
<td>.24</td>
<td>.18</td>
</tr>
<tr>
<td>Counting</td>
<td>.500</td>
<td>.08</td>
<td>.01</td>
<td>.05</td>
</tr>
<tr>
<td>Displaced Rehearsal</td>
<td>.499</td>
<td>.04</td>
<td>.14</td>
<td>.05</td>
</tr>
</tbody>
</table>
The failure to demonstrate any difference between uncontrolled and controlled response order conditions is interesting in its own right. Subjects in the former condition do tend to fill in their responses for items from the recency and primacy portions of the list first, and work their way in (this observation is based on the experimenter's observations of over 400 subjects). Subjects in the latter condition were unable to do this, but performance did not appear to be adversely affected.

**Protocol Analysis.** The strategy protocol data shown in Table 3 were collected on the initial hypothesis that the free-recall training advantage in the first experiment, might have resulted from an increased tendency for subjects in this condition to engage in elaborative strategies, and/or a tendency for subjects in the serial-position judgment training condition to use inappropriate strategies such as counting, or assigning a number to represent each item's position in the series. Also, these protocols might provide a test for the encoding assumptions outlined above. Subjects were instructed to describe the strategy they used while studying the Trial 2 list, and to indicate whether this was the same strategy they had used on the Trial 1 study-list).

Protocols were collected from 271 subjects. An initial analysis of these protocols was made by a single judge. The judge was guided by classifications of strategy types commonly accepted in the relevant memory literature (Weinstein et al., 1979; McDaniel & Kearney, 1984; Shaughnessy, 1981). From this analysis eight types of strategies
were identified (at least 3% of the subjects reported using each one). Following this initial screening three judges (the original judge and two others) examined these initial classifications. For a given protocol to be included in one of these categories all three judges had to agree that it belonged there. The eight strategy types, and the defining characteristics of each were as follows:

1) Story- The subject reported trying to make up a story or stories from the words.

2) Complex Visual Imagery- The subject reported that they created interactive images from two or more adjacent items.

3) Chunking- The subject reported trying to rehearse groups of adjacent items (ranging in size from 3 to 10 items).

4) Rote- The subject reported simply rehearsing the current item until it went off the screen.

5) Counting- The subject reported assigning a number to represent an item's position in the series.

6) Displaced Rehearsal- The subject reported rehearsing items that were related, but not adjacent in the list (for example, rehearsing items that are associatively related, semantically related, or share similar phonological structure).

7) Mixed- The subject reported employing more than one type of strategy on the same list.

8) Not Categorizable- If the subjects reported strategy was not clearly verbalized or was very idiosyncratic, they were placed in this category.
There are two issues that are addressed by these data. First, the relative efficacy of different types of chosen strategies in a temporal-order judgment task can be determined. Second, the degree to which prior experience with the task induced subjects to shift strategies, and the extent to which these chosen strategies differed from the incidental group, also can be determined. These two issues will be discussed separately in the following sections.

Effects of chosen strategies. To the issue of whether chosen strategies influenced performance, the data indicate, quite clearly, that they did. Further, the pattern of data conforms to the predictions derived from the current model. Recall that it was argued that localized rehearsal patterns among temporal neighbors is the necessary condition for temporal-order learning. While both chaining and chunking processes would accomplish this, chunking might be superior in that items might be indirectly linked with more temporal neighbors than if simple chaining was employed. Further, the chunk-level codes would provide a second level of ordinal structure with fewer links to be maintained, and if subjects access these codes, then reasonably accurate position judgments could be made. In any event, either chaining or chunking should be superior to displaced rehearsal patterns (which should inhibit localized rehearsal patterns).

Interestingly, only two subjects reported simple chaining strategies, but 20% of the subjects reported simple chunking strategies (chunks ranged between three and ten items, but reports of three through five were the most common, also a number of subjects did
not specify the group size). The mean z score for this group is .690. Comparing this condition with the displaced rehearsal cell, with a mean Z score of .499 (note that this group had the lowest score along with the counting group with a mean of .500), the difference was significant $t(70) = 2.4$, $p = .008$ (probabilities of one-tailed t tests are reported for all these data). Therefore, simple chunking was superior to displaced rehearsal patterns, and this supports quite strongly the central assumption of the current model; namely, that localized rehearsal is critical in order learning.

A comparison of simple chunking with counting strategies (reported by 5% of the subjects) indicates that chunking was the superior strategy, $t (28) = 1.9$, $p = .03$. That outcome is consistent with the idea that chunking should be superior because multiple links are created. After all, chunking produces several temporally relevant links, whereas counting produces only one relevant link (with the number itself). Assuming a certain amount of forgetting, chunking should be the more efficacious strategy of the two. Obviously, forgetting must have been the cause of this disadvantage for the counting strategy, since if the number was remembered, performance in the counting group would be perfect. In short, the advantage for simple chunking over both displaced rehearsal and counting provides support for the crucial assumptions of the current model. Both counting and displaced rehearsal strategies would inhibit the process of relational coding of an item with its temporal neighbors.
The complex visual imagery, and story strategies produced the best performance (mean z scores of .890 and .947 respectively). The difference between these scores was not significant, $t(20)=1.6$, $p=.06$. The difference between imagery and simple chunking strategies was not significant $t(68)=1.5$, $p=.06$. However, the difference between stories and simple chunking was significant $t(111)=2.9$, $p=.004$. Further, with the exception of the failure to obtain a difference between the imagery and simple chunking groups, the story and imagery strategies were superior to all other types.

The visual imagery group contained subjects who reported creating complex images for pairs, or groups, of adjacent items. Thus, the condition of localized rehearsal, as well as chunking, would be met by this elaborative strategy. Generating stories from the words would also induce both localized rehearsal patterns and chunking (each item would be added to the story structure based on its position in the input series, and therefore would be linked with more than one other item, in creating coherent scenes or episodes). That these three strategies resulted in the three highest scores, can be taken as strong support for the assumptions of the current model. All three strategies were superior to counting or displaced rehearsal.

Subjects who reported using more than one strategy within the list (mixed), did fairly well (mean z score=.638). This strategy was significantly better than displaced rehearsal, $t(50)=2.1$, $p=.01$, but the difference between mixed and counting was not significant, $t(60)=1.5$, $p=.06$. However, mixing strategies resulted in
significantly poorer performance than either visual imagery, t(60)=2.7, p=.003, or stories, t(93)=4.0, p=.0001. However, the difference between simple chunking and mixed was not significant, t(95)=.72, p=.46.

Since the types of strategies employed by subjects in the mixed group varied a good deal, any interpretations of these results are speculative. However, as noted earlier, the use of more than one strategy across the list might create meaningful changes in process context (Block, 1982; Underwood, 1977), and that information could, in principle mediate a serial-position judgment. For example, remembering that strategy A was used first, and B second, would tend to make different sections of the list distinctive. If when judging the position of a target the subject could remember how it was encoded, then its position in the series could be approximated. This mixing of strategies therefore, might be superior to uniformly bad strategies such as displaced rehearsals, and the data indicate that this was the case. However, since varying strategies does not insure continued localized rehearsal patterns, it might be inferior to chunking or elaborative chunking strategies in general, and as noted the mixed group performed worse than either the imagery or story groups.

The one unanticipated result concerns the relatively good performance of subjects who reported engaging in rote, or maintenance rehearsal (mean=.661). Only 3% of the subjects reported using this strategy (note that that seems inconsistent with the claim made by
authors like Shaughnessey (1981), that rote rehearsal is, for most subjects, a preferred strategy. This type of strategy resulted in significantly poorer performance than either the imagery, $t(20)=1.7$, $p=.055$, or story strategies, $t(17)=2.6$, $p=.009$. However, no other differences were reliable.

It might be speculated that subjects in this group, by not consciously engaging in either localized or displaced rehearsal patterns, will automatically link the current item with temporal neighbors still active in working memory, and that would imply at least some localized rehearsal. That would not, however, imply that temporal-order learning is automatic, since consciously chosen rehearsal patterns can either facilitate or interfere with this learning, relative to this rote condition. The problem with this sort of speculation is that statistical analyses indicated that the apparent advantage for the rote group over the displaced rehearsal group, or the counting group, was not reliable. The fact that so few subjects actually reported using this strategy might account for this lack of reliability. In any event, no firm conclusions concerning the relative efficacy of rote rehearsal can be drawn from this data.

In summary, there is clear evidence from this data that subjects chosen strategies are correlated with performance in a temporal judgment task. This finding provides the first clear documentation of this effect to date, and strongly argues against the claim that temporal-order learning is an automatic process (one that is uninfluenced by controlled strategic processing).
Further, the pattern of data conforms quite well to the assumptions of the current model, in that strategies that induced localized patterns of rehearsal (and chunking or grouping), were superior to strategies that induced either displaced rehearsal patterns, or the establishment of single as opposed to multiple temporally relevant links. The simple chunking, imagery, and story strategies were all superior to both, displaced rehearsal, and counting strategies, and these latter two strategies should inhibit associative linking with temporal neighbors.

It is not clear how any alternative assumptions could account for these data. For instance, one might argue that perhaps the performance differences were simply the result of differences in level of processing for the items in these different strategies. In general, items that are processed semantically are better recalled than items processed in a more shallow fashion (such as when subjects attend to spelling patterns or other non-semantic properties of items) (Craik & Lockhart, 1972). Thus, the advantage for the imagery and story conditions might occur because in these conditions words were more likely to be processed semantically than in other conditions, and therefore less likely to be forgotten.

There are two reasons why this account is probably faulty. First, most of the subjects in the displaced rehearsal condition reported grouping non-adjacent items that were from the same semantic or associative categories, and this implies semantic level processing. Subjects in the simple chunking group, on the other hand, reported
simply rehearsing groups of adjacent items, and this should be a far less elaborative or semantic strategy, since the item's semantic or associative category is not being considered. Free-recall data indicate that categorical clustering produces better recall, in general, than simple rehearsal processes (Crowder, 1976). The fact that subjects in the chunking group did better than those in the displaced rehearsal group, then, would argue against a simple levels-of-processing interpretation of this pattern of data.

Second, a number of studies using similar materials and list lengths have reported near ceiling recognition performance (Zacks, et al., 1984; Michon & Jackson, 1984; Tzeng, 1976). Tzeng (1976), for example, found that recognition performance with identical list lengths, and concrete nouns, did not differ between subjects engaged in categorical clustering and those engaged in rote rehearsal; both groups, in fact, were near ceiling performance (although the clustering group did better on a recall test). Therefore, the simple forgetting account implied in the levels of processing argument, would not seem to be tenable (remember that a serial-position judgment task uses a recognition, as opposed to a recall format).

Of course, the key problem with this self-report data is that it is merely correlational evidence. That is, it cannot be determined whether the choice of strategies caused the performance differences, or were simply correlated with other factors that might have produced differences (such as subjects general intelligence, for example). The third experiment in this series addresses this concern.
Effects of training on choice of strategies. Given that chosen strategies appear to have had an influence on performance in a serial-position judgment task, which violates the critical assumption of the automaticity hypothesis, it can be asked why it is the case that both manipulations of instructions, and prior experience, fail to produce overall performance differences. As noted earlier, it may be that subjects in the expectancy conditions simply do not know what types of strategies are more efficacious for order learning. Alternatively, it may be the case that subjects expecting an order test do employ generally good strategies, but that subjects in the incidental conditions also tend to employ similar strategies for different reasons. It would be interesting to know whether the majority of subjects in both conditions are employing relatively good, or relatively poor strategies, and an analysis of chosen strategies as a function of training condition might address these issues.

It might be the case that the current training manipulation was not adequate to induce shifts in the patterns of chosen strategies, and that could account for the failure to obtain an overall performance difference between the two groups. However, in the current experiment 51% of the subjects in the serial-position training condition reported using different strategies on the second trial than on the first trial, as compared with 35% of those subjects trained with a free-recall task, and this difference was significant $X^2 (1) = 5.24, p<.02$. Thus, the likelihood of shifting strategies as a function of training was greater for those subjects trained with a
temporal judgment task. Further, an analysis of strategy type (eight types) by training condition (two types) revealed that the types of strategies chosen differed between the two groups, \( \chi^2 (8) = 28.0, p = .000 \). Therefore, it can be concluded that the current manipulation did induce subjects trained with a serial-position judgment task to use somewhat different strategies than those subjects in the incidental condition (free-recall training).

An examination of Table 3 reveals an interesting pattern to this data. First, while 51% of the subjects in the serial-position training group used one of the three best strategies (chunking, visual imagery and stories), 40% of the subjects in the free-recall training groups reported using these as well. An analysis revealed that the probability of choosing any one of these three strategies as a function of training was not significantly different \( \chi^2 (2) = 3.5, p = .16 \). A fairly large percentage of subjects chose good strategies (one that induced localized rehearsal patterns and chunking), whether primed to expect a temporal judgment task or primed to expect a free-recall type of test. It might be concluded therefore, that the incidental condition, in general, does not necessarily induce subjects to employ poor strategies, and that might account for the failure to find intentionality effects in the temporal coding literature.

Consider next the two poorest strategy types (counting and displaced rehearsal). While 8% of the subjects in the serial-position training condition chose the counting strategy, only 1% of the subjects trained with free-recall made that choice. Conversely, 14%
of the subjects in free-recall training chose displaced rehearsal strategies, compared with 4% in the serial-position group. An analysis of the likelihood of choosing one of these two strategies as a function of training was conducted and a significant difference was obtained, $\chi^2(1)=15.5$, $p=.000$. Thus, subjects trained with free-recall were more likely to displace rehearsals, and subjects trained with serial-position judgment tasks were more likely to choose counting strategies. However, the numbers of subjects employing these two strategies are very close (5% counting vs. 8% displaced rehearsal). Therefore while subjects in the incidental condition were somewhat more likely to displace rehearsals, this did not produce an overall disadvantage for this group since an almost equal number of subjects primed to expect an order task, also chose an inappropriate strategy (counting).

In summary, given that subjects in the two groups were approximately equally likely to choose good or poor strategies, although there are differences in the patterns of choice, it is hardly surprising that no overall difference between the two conditions was observed. It certainly does not imply that processing is automatic, as had been previously assumed, given the significant differences in performance as a function of strategy. Subjects in both conditions tended, by and large, to employ relatively good strategies in coping with the demands of the experiment, which suggests that many subjects have a fairly well developed knowledge of the kinds of strategic processes that result in good memory performance (good metamemory).
Experiment 3

The implication from the protocol data seems to be that certain types of reported strategies can facilitate the representation of temporal-order information. Apparently, strategies that induce localized as opposed to displaced patterns of rehearsals, produce better temporal-order learning. Strategies that do not meet these conditions, such as displaced rehearsal or counting strategies, produce relatively poor performance. That is consistent with the predictions of the current model. However, a controlled experiment is needed to establish that strategies cause performance differences, and that they are not merely correlated with them. In order to accomplish that goal the current experiment manipulated strategies as an independent variable.

In this experiment subjects were assigned to one of three strategy conditions: pleasantness rating, which is highly item-specific (Hunt & Einstein, 1981); general relational, which should induce displaced rehearsal patterns; or constructing stories to string the words together, which should induce both chunking and localized patterns of rehearsal. Performance on a serial-position judgment task was, once again, the dependent measure in the experiment.

The story strategy resulted in the best performance in Experiment 2, and satisfies both the chunking and localized rehearsal conditions (see above), and should induce a considerable amount of relational processing among items (items are related within the context of the story episodes or scenes). This can be directly compared to a highly
item-specific strategy (pleasantness rating), which should minimize relational coding. The pleasantness-rating task was chosen over a simple rote rehearsal condition, for the following reasons. First, the pleasantness rating task insures that items will be processed at a semantic level, and this avoids a levels of processing confound. Second, the data from the Experiment 2 suggested that perhaps rote rehearsal, which is not a very attentionally-demanding cognitive task, may have allowed a certain amount of implicit relational processing. The more demanding pleasantness rating task, might more effectively limit associative coding with temporal neighbors, and would therefore be a better contrast with the story strategy.

In the general relational coding condition, subjects were instructed to look for ways in which a current item was related to, or similar to, earlier items from the list. This strategy should induce a high degree of displaced rehearsal (since associative relationships were uncorrelated with input order). Therefore, a comparison of localized and displaced rehearsal is afforded when the story and general relational strategies are compared (this will allow for a replication of the story and displaced rehearsal performance differences observed in Experiment 2, to be attempted).

A comparison of the general relational, and pleasantness rating strategies will also make it possible to determine whether displaced rehearsal patterns add any information, above and beyond item-specific coding. Displaced rehearsal theory would argue that the associations formed between non-adjacent items preserve both relative order, and
lag, or distance, information. On that account it would be anticipated that such a strategy would be superior to the highly item-specific processing induced by the pleasantness-rating strategy.

Method

Design. A single factor between subjects design (with three levels of this variable) was used. The variable was strategy instructions, and the three groups were: pleasantness rating, relational, and stories.

Materials. The materials were the same as in the previous experiments. Since the type of test (controlled vs. uncontrolled order of responding) did not have any affect in the strategy protocol data from Experiment 2, and no effect in general, all subjects were given the test sheets with all the items listed in a column. Also, in this experiment, each subject was given only one trial, so training influences were not a factor.

Subjects. The subjects were 61 Psychology students drawn from the same population as the other experiments.

Procedure. Subjects were tested in groups of between four and eight. All subjects were informed that the experimenter was interested in determining the effects of strategies on memory for items, and that they should try to use only the strategy assigned to them, and not deviate from it. Each group was told that their particular strategy had been shown to facilitate memory. In the pleasantness rating condition subjects were told that they would be shown a series of words, and when they saw a word they should write
down on a piece of paper the letter P if the word appeared to them as pleasant, or a U if unpleasant. They were instructed to restrict themselves exclusively to this task. Subjects in the relational group were told to think about ways in which the word currently on the screen was related to other words from the list (types of relationships were not specified). Finally, subjects in the story condition were told to try to make up a story, or stories, from the words. Each group was then shown the study list, and subsequently performed the serial-position judgment task. Again, all other conditions were the same as in the first two experiments.

Table 4: Data from Experiment 3

<table>
<thead>
<tr>
<th>Condition</th>
<th>Type of Strategy</th>
<th>Mean Z* Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Story</td>
<td>.977</td>
<td></td>
</tr>
<tr>
<td>Relational</td>
<td>.639</td>
<td></td>
</tr>
<tr>
<td>Pleasantness</td>
<td>.627</td>
<td></td>
</tr>
</tbody>
</table>

\[ \bar{X} = .748 \]
Results

The data from this study are shown in Table 4. The strategic variable had a significant effect on performance $F(2, 58) = 4.69$, $p < .01$, indicating that performance in the story condition was better than in the pleasantness-rating or relational conditions. Subsequent analyses confirmed that the difference between the story condition and the relational condition was significant $F(1,39) = 5.32$, $p < .05$. The difference between the story and pleasantness rating conditions was also significant $F(1,39) = 6.02$, $p < .05$, but the difference between the relational and pleasantness rating conditions was not significant, $F < 1.00$.

Discussion

These data quite clearly support the hypothesis that strategic variables can affect temporal judgment performance, even when strategies are manipulated by the experimenter, instead of simply being chosen by the subject. Therefore, it can be stated more confidently that strategies produce performance differences, and are not merely correlated with them.

As in Experiment 2, the story strategy (which satisfies the localized rehearsal and chunking conditions) resulted in better performance than the general relational strategy (which should result in a good deal of displaced rehearsal). Further, the pleasantness-rating task, which limits relational coding in general, also resulted in poorer performance than the story strategy. Therefore, the general advantage for the story strategy replicates the pattern of data
obtained in Experiment 2, and provides further support for the assumptions outlined in the current model.

General relational coding was not significantly better than pleasantness-rating. Apparently, both displaced rehearsal patterns, and the item-specific processing induced by the pleasantness-rating task, interfered with the localized rehearsal patterns necessary for good temporal-order learning. Further, the displaced rehearsals apparently added no information pertinent to the serial-position judgment task. Thus, displaced rehearsal does not appear to provide any useful information for temporal-order learning. In fact the data from both of these experiments indicate that it actually inhibits this type of learning (this is also consistent with the data reported by Pugh (1986) that categorical clustering can also adversely affect paired-recency performance for items not rehearsed together).

Summary of Experiments 1-3

In Experiment 1, the training manipulation appeared to result in a slight advantage for subjects trained with free-recall over those trained with a serial-position judgment task. The data from Experiment 2 indicated that this advantage was not reliable. However, the protocol data collected in this experiment revealed that chosen strategy did appear to affect temporal-judgment performance. Strategies that induced localized patterns of rehearsal, and chunking or grouping were superior to strategies that induced either, displaced rehearsals, or single temporally relevant links (counting strategies).
That pattern of data conforms to the predictions of the current model (namely that localized rehearsal patterns are necessary conditions for temporal-order learning).

Further, the failure to obtain differences as a function of expectancy seems to have been a consequence of the fact that subjects in both training conditions were equally likely to choose good or bad strategies, although the specific pattern of choices differed somewhat. The critical point is that a majority of subjects in both the incidental and expectancy conditions chose relatively good strategies, in general. Finally, the third experiment is consistent with the protocol data, and indicates that these performance differences are not merely correlated with strategy, but rather appear to be directly determined by the strategies themselves.

Clearly, the current findings strongly argue against the automaticity hypothesis as stated by Hasher and Zacks (1979). The current experiments provide the first direct evidence that chosen strategies influence performance, and also indicate why expectancy effects are not, in general, obtained.

Also, it should be mentioned that these strategic effects, while consistent with the relational coding model, would strongly argue against the notion of an invariant time tag proposed in time tag models. On these accounts, there would be no reason to predict differences in performance as a function of encoding strategy. The data suggest that future research should focus on patterns of associative rehearsal, in order to account for temporal-judgment
performance. Treating temporal coding as a special process, as implied in both the automaticity hypothesis, and time tag models as well, would appear to be unwarranted.
As noted in the section on retrieval and decision assumptions, there is a good deal of evidence that temporal position information can be forgotten fairly rapidly (Underwood, 1977). If there is little, or no, direct temporal information available at the judgment stage, will subjects simply guess, or will their responses be mediated by the perceived strength of the memory-trace, which subjects might reasonably assume decreases with time?

The notion that memory-trace strength might serve as a cue to recency has a long history in experimental psychology (Hinrichs, 1970; Michon & Jackson, 1984). The idea is that subjects will assume that the weaker the memory-trace strength, the further back in time (or the list) the item must have occurred. While it is clear from the preceding experiments that subjects rely primarily on associative information in making judgments, the question of whether, under certain conditions, subjects will be influenced by memory-trace strength, remains an important one. If it is determined that subjects do exhibit a strength bias under certain conditions, then the role this bias might play in any given experiment would need to be considered, in order to provide an adequate account of performance.
The general aim of the current experiments was to create conditions where direct temporal information would be minimally available, and manipulate certain variables that should affect memory-trace strength, to determine whether these manipulations influence performance.

In a pilot study the number of exemplars from a given taxonomic category within a long study list (104 nouns) was varied. Some categories had three exemplars, some six, and some nine exemplars within the list. After the study phase a paired-recency test was given to subjects containing 48 test pairs. Of these, 24 pairs contained items from the same categories (apple-pear) while 24 contained items from different categories (apple-vodka). Of this latter type, the items in a test pair came from categories of sizes: 9 and 6, 9 and 3, or 6 and 3 (based on number of exemplars in the list). Within each of these, an equal number of pairs had an item from the larger category that was less recent, or more recent.

Certain models of recognition memory predict that an item from the category with more exemplars in the list will have a higher memory-trace strength value (because of associative links), and will be subjectively more familiar (Mandler, 1980; Gillund & Shiffrin, 1984; Hintzman, 1988; Rumelhart & McClelland, 1986). Consistent with this notion the likelihood of hits and false alarms (in other words the likelihood of saying yes) in recognition judgments are greater for items from categories with a larger number of exemplars in the study list, than for items from categories with fewer exemplars (Mandler &
Rabinowitz, 1981). The memory-trace strength for the item from the larger category is greater on these trace-strength accounts. (It should be noted that strength or subjective familiarity accounts are by no means the only plausible ones, however, the current experiments are predictated on the notion that memory-trace strength is a viable theoretical construct).

If paired-recency judgments are mediated by memory-trace strength, then when the item from the larger category is actually more recent than the item from the smaller category, a strength bias should result in increased accuracy. Conversely, when the item from the larger category is actually less recent, the strength bias should lead to decreased accuracy. That pattern would be most likely to be obtained under conditions where direct temporal information is minimized (where overall performance is near chance levels).

In order to induce poor direct temporal coding in this experiment, a long list was used, with fairly rapid presentation rates (2 seconds per item). Also, test pair items never came from the beginning or end of the list (to avoid primacy or recency effects), and the lag between the target items was short (mean of 5). In fact, performance hovered at around 50% correct (chance) in these data. However, no differences were obtained as a function of category size for pairs from different categories. In short, no strength bias as a function of this independent variable was observed.

However, one interesting finding emerged when items from the same category were compared. When pairs were composed of items from
categories with nine exemplars, performance was worse than when they were from categories of sizes six and three (pooled for this analysis to give an equal number of pairs for the analysis). It may be that items from categories with more exemplars were more likely to be clustered, producing less localized rehearsal for these items, and that this inhibited the establishment of temporal position information for these items. That would seem to make sense in light of the apparent categorical clustering disadvantage reported by Pugh and Johnson (1987), as well as the data from Experiments 2 and 3 of the current study.

Experiment 4 manipulated word frequency in order to determine whether this manipulation, which might affect the memory-trace strength of items, will produce evidence of a strength bias in temporal judgments. Experiment 5 used a priming paradigm to manipulate trace-strength.

Experiment 4

In this experiment a paired-recency judgment task was used to examine the effects of word frequency on temporal judgment performance. In general, high frequency words are more accurately recalled than low frequency words, but the opposite pattern obtains in a recognition task (Lachman, Lachman, & Butterfield, 1979; Mandler, 1979; Schulman & Lovelace, 1970). The advantage for low frequency words in recognition tasks has received many accounts in the literature, but the strength account assumes that the presentation of
a low frequency word in the study-list increments its memory-trace activation level, relative to its normal state, to a greater degree than is the case for high frequency words (Mandler, 1979; Gillund & Shiffrin, 1984).

On that account one might expect that the low frequency items will seem relatively more familiar than normal after presentation in a study list. This relatively large increment in trace strength might lead to a tendency to judge these items as having occurred more recently in a temporal judgment task. Specifically, if a test pair in a paired-recency judgment task contains a low frequency item and a high frequency item, and if the availability of other, more direct temporal information is minimized, performance should be quite poor when the low frequency item is actually less recent than the high frequency item, but well above chance if the low frequency item is actually the more recent item.

Method

Design. A single variable was manipulated in this study: type of test pair (with four levels) in a within subjects design.

Materials. A single study list was constructed containing 84 words. Of these 20 were fillers (10 from the beginning and 10 from the end of the list). Test pairs were constructed from 64 items for the paired-recency test. Of these, 32 were low frequency words (mean Kucera and Francis (1968) value = 3.53), and 32 were high frequency (mean Kucera and Francis value = 267.4). Words varied in length from one to three syllables.
Four types of test pairs were constructed from these 64 words. Eight pairs contained two high frequency words (high-high), eight contained two low frequency words (low-low), eight contained a low frequency word as the less recent item and a high frequency word as the more recent item (low-high), and there were eight others with the opposite arrangement of high and low frequency items (high-low). For each of these four types, the mean lag between the items, and mean lag of the more recent item, were equated (see Appendix A).

The study list was presented on a Televideo 920C terminal at a rate of 1500 msec per item (this rate was chosen to minimize relational coding processes. The test pairs were also presented to subjects on the terminal, one above the other. Subjects pressed the 'Z' key on the keyboard if they thought the item shown at the top was more recent and the '?' key if they thought the item on the bottom was more recent. Presentation order for test pairs, as well as which member appeared on top or bottom, was randomized between subjects.

Subjects. The subjects were 74 introductory Psychology students from the same population as the above experiments.

Procedure. Subjects were run individually. They were told that they would be studying a list of words to be followed by a memory test (the test was not specified). During a two minute retention interval subjects were told that they would see two words from the study list, and that they should choose the one that occurred more recently than the other one, by pressing the appropriate response key.
Results

Mean percent correct in each of the four test pair conditions was the dependent measure and transformed using the arcsine transformation procedure.

The data for each of the four conditions are presented in Table 5. Analyses revealed an effect of test-pair-type \( F(3, 219) = 5.12 \) \( p = 0.001 \). The data indicate that while performance hovered around 50% correct in general, performance in the low-high condition was quite poor (43%). Separate analyses revealed that this condition differed significantly from each of the others (low-high vs. high-low, \( F(1, 73) = 5.66 \), \( p = 0.02 \); low-high vs. low-low, \( F(1, 73) = 13.75 \), \( p = 0.000 \); low-high vs. high-high, \( F(1, 73) = 8.19 \), \( p = 0.005 \)). However, the other three were not significantly different from each other (all \( F \)'s < 1.0).

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean % Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-High</td>
<td>54</td>
</tr>
<tr>
<td>High-Low</td>
<td>51</td>
</tr>
<tr>
<td>Low-High</td>
<td>43</td>
</tr>
<tr>
<td>Low-Low</td>
<td>54</td>
</tr>
</tbody>
</table>
Discussion

Overall, a tendency for low frequency items to be judged as more recent when compared with high frequency items was obtained, suggesting that a strength bias may be operating. However, the failure to obtain an advantage when the low frequency item was actually more recent is not consistent with this account. It should be noted that Wolff (1968) in an attempt to manipulate memory-trace strength by comparing paired-recency performance on pairs consisting of both low and high associative value CVCs obtained a very similar pattern of data (see above). However, the reason for this pattern of data in both studies is entirely unclear.

It cannot be concluded from this experiment that a strength bias was operating in these conditions. While the disadvantage for low-high pairs is consistent with this idea, the failure to obtain a corresponding advantage for the high-low pairs is not. Given that other factors such as lag and frequency were equated for the four conditions, there is no reason, under the strength bias hypothesis, to expect that the high-low pairs would not be judged more accurately than the two control conditions (high-high, and low-low). If a strength bias was operating at all in this experiment, its presence is not unequivocally demonstrated in the pattern of data.

It might be argued that the failure to obtain an advantage for high-low pairs could have come about because of some experimental artifact in this cell. For example, while the frequencies and lags in this condition were equivalent to other cells (see Appendix A),
perhaps some of these items had lower or higher subjective frequencies than the norms would suggest. The problem with this type of argument is that Wolff (1968) using high vs. low associative value nonsense syllables obtained exactly the same asymmetrical effect. Therefore, if there is anything systematic about this finding, the artifact explanation seems unlikely. Since the asymmetry is inconsistent with the memory-trace strength account, it might be necessary to look elsewhere for an explanation of the poor performance in the low-high cell. In any event, with regard to the issue of whether memory-trace strength biases temporal judgments, the current data do not permit a positive conclusion.

Experiment 5

In this experiment subjects were shown a list of eighty words in the study stage. Subsequently, they were given a test list that contained these items along with a number of new items (foils). They were to decide if: 1) The test item was from the study list (recognition task), and 2) if they thought that it had been in the study list, they were to assign a number from 1-8 that represented in which eighth of the list the item occurred (serial-position judgment). The critical manipulation in this experiment concerned the foils (new items). Some of these foils were strong associates of items that had actually been in the study list (e.g. if DOCTOR was in the study list, NURSE might be shown as a foil in the test stage). These foils occurred in the test list either immediately before the associated
target (lag=0), separated by one intervening test item (lag=1), or separated by between 2 and 6 test items (lag>1). These three conditions can be compared with a control condition where the related foil was not present in the test list.

The hypothesis is that in processing the foil, the related target will be automatically activated (primed), and when the target is subsequently presented for a position judgment, this increased activation received during the test stage, will make the item seem to have occurred more recently than it actually did (assuming an increment in trace strength). This should result in primed items being assigned to a later portions of the list more often than unprimed items (therefore the deviation from true position in the judgments should be more systematically in the direction of the later portions of the list for primed items, than for unprimed items).

Ratcliff and McKoon (1981) reported that in a recognition judgment task, judgments for target items are made more quickly, if they are preceded in the list by a strong associate. On the memory-trace strength hypothesis this would result from the fact that the priming incremented the strength value of the target making it easier to access. While this account is by no means the only one possible, since it is the trace-strength hypothesis as a temporal bias that is being explored in the current experiments, a priming manipulation would seem to be an appropriate test.
Method

Design. A within subjects design was used in this experiment. The critical manipulation in the position judgment task was priming. The four conditions were: unprimed, lag=0, lag=1, or lag>1.

Materials. A single study list containing eighty words was generated. Of these, 32 words that possessed strong associates were chosen using norms reported by Keppel and Strand (1977). These words were distributed evenly from position 11 through 70 in the study-list (the second through the seventh eighth of the list (each eighth of the list consists of 10 items).

Four different test lists were composed according to the following scheme. For every target item its associate was not a foil in one of the lists (unprimed), or its prime occurred at one of the three lags in the other three test lists. In this way each target item appears in each of the priming conditions across subjects (thus each item serves as its own control in the unprimed condition).

The study and test lists were presented to subjects on a Televideo 920C terminal. The study list was presented at a rate of 4 seconds per item. This somewhat slower rate might improve relational coding compared to the previous frequency study (where 1.5 second rates were used), but it was deemed necessary to allow for sufficient encoding time so that recognition memory would be accurate enough to make it possible for subjects to distinguish targets from foils.

Test items were presented to subjects one at a time, and they were to press the 0 key if they thought the item had not been in the study
list, or a key between 1 and 8 to estimate that eighth of the list that a recognized item had occurred in.

Subjects. The 74 subjects from Experiment 4 served in this experiment as well. Each subject participated in both experiments within the same session, but order of experiments was counterbalanced to avoid any set effects.

Procedure. Subjects were run individually, and were told that they would be presented with a study-list, and that after that, a memory test would be given (the nature of the test was not specified, although half of the subjects had already been tested with a paired-recency task in Experiment 4). Prior to the test stage subjects were instructed regarding the serial-position judgment task, and instructed to use the keys 0-8 to indicate their judgments (the retention interval was approximately 2 minutes).

Results and Discussion

For every judgment in the four priming conditions a signed deviation score was computed. For example if the item occurred in the fifth segment of the list a judgment of 5 received a score of 0, a judgment of 4 a score of -1, or a judgment of 6 a score of +1, and so on. The mean signed deviation score four each condition is shown in Table 6.
Table 5: Mean deviation scores for priming conditions in Experiment 5

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean Deviation score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unprimed</td>
<td>+1.5</td>
</tr>
<tr>
<td>Lag=0</td>
<td>+1.9</td>
</tr>
<tr>
<td>Lag=1</td>
<td>+2.6</td>
</tr>
<tr>
<td>Lag&gt;1</td>
<td>+2.3</td>
</tr>
</tbody>
</table>

An analysis of variance indicated that there was no effect of priming condition on performance in this experiment $F<1.00$. Therefore, it can be concluded that the priming manipulation, which should produce differences in memory-trace strength between primed and unprimed items, had no influence on the performance in the serial-position judgment task. Thus, once again the strength-bias hypothesis was not supported.

Summary of Strength Experiments

With three independent manipulations that have been thought to affect memory-trace strength (number of categorical exemplars, word frequency, and priming), no clearcut evidence of a strength-bias in temporal-judgment performance was obtained. These three studies converge with evidence reported by Wells (1974), wherein recognition confidence ratings were uncorrelated with performance on an absolute judgment of recency task. Together, these studies strongly question that claim that memory-trace strength is a factor in temporal judgment performance. While each of the four studies may be flawed in some way,
if memory-trace strength is a viable cue to temporal position, it would be expected that some evidence would have emerged in at least one of these studies (Garner, Hakes, & Ericksen, 1956).

At present, therefore, it can be concluded that there is no support for Assumption 4 of the current model from these four manipulations. If subjects do not rely on memory-trace strength as a source of temporal information, then a sufficient account of performance is contained in the encoding assumptions outlined above. Certainly, it would be unwarranted at present to complicate the issue with a consideration of response bias.
CHAPTER VI

CONCLUSIONS

The current set of five experiments were conducted to address several related issues pertaining to temporal-order learning. Four assumptions were proposed initially, and the experiments produced a number of results that are relevant to these ideas. These four assumptions, and the extent to which the current results are consistent, or inconsistent, with them, will be discussed in this section.

Assumption 1

Assumption 1 proposed that there is no invariant time tag that specifies an event's position in time. Time tag models claim that one of the attributes that will be contained in the memory representation for an item, represents, on at least an interval-level scale, when that event occurred in time. Biological clock models, as well as spatial analog models represent this class of theories (Treisman, 1963; Bindra, 1976; Murdock, 1974; Moyer & Bayer, 1976). These types of models all predict that increasing the temporal interval between stimulus items will improve performance in a temporal judgment task. Previous manipulations of inter-stimulus-interval (ISI) have, in
general, failed to support this prediction (Greene, & Bower, 1988; Guenther & Linton, 1975; Shiffrin & Cook, 1978).

The current findings are also inconsistent with the time tag notion. The fact that the kinds of rehearsal patterns that subjects employed, affected temporal-order learning, suggests that temporal position information is contained in the associations formed between a given item and its temporal neighbors (relational coding processes). According to time tag models temporal position is established, not with reference to other items, but instead with reference to some type of abstract time scale. On that account relational coding processes should have little effect on performance.

If it were assumed that time tagging requires a certain amount of attentional effort, then it could be argued that displaced rehearsal strategies might interfere with this process in the following way. Displaced rehearsals between the current item, and earlier related items might discourage subjects from attending to the temporal position of the current item. If establishing a time tag requires that subjects consciously attend to the current temporal context (or value of some type of biological clock), then the negative influence of displaced rehearsal strategies could be explained. However, on that account, counting strategies, which do encourage subjects to attend to an item's position in time, should facilitate, or at least not interfere with, time tagging. The data indicate that counting strategies and displaced rehearsal strategies produced equally poor performance on the serial-position judgment task. On the other hand,
strategies that encouraged localized relational coding processes (such as chunking) resulted in superior performance. There is no reason on the time-tag account to assume that these relational coding processes would improve performance.

However, that pattern of data is consistent with the claim that temporal position information for a given item is preserved through links with temporal neighbors, and this is inconsistent with the claim that true interval level temporal information is represented in this context. Therefore, Assumption 1 appears to have received support in these experiments.

Assumption 2

Assumption 2 proposed that, in the absence of temporally relevant extra-list context, the only information that would specify an item's temporal position would be the associative links formed with temporal neighbors. This, of course, is predicated on the assumption that there is no invariant time tag. Either serial-chaining (wherein an item is linked directly with items that immediately precede or follow it), or chunking (wherein an item is linked through the chunk code with several temporal neighbors), would satisfy this condition for order learning. If an item is not linked with its temporal neighbors, then a temporally-ordered representation of the stimulus list will not be established. Without special time tags, or extra-list markers, the process of temporal coding reduces to one of serial-order learning.

Further, given that encoding strategies have been shown to influence serial-order learning when measured by serial-recall (or
serial-anticipation learning), it would seem to follow that temporal coding is not an automatic, or mandatory, process. Therefore, it was argued that different types of cognitively-controlled encoding strategies will be differently effective in temporal-order learning.

Strategies that induce localized rehearsal patterns, such as chaining or chunking (including elaborative chunking strategies), should result in better temporal judgment performance than strategies that induce displaced rehearsal patterns. Further, chunking strategies might be superior to simple chaining strategies, for the following reasons. First, the chunk codes create a second level of organization, which subjects might refer to when making temporal judgments. Since there would be fewer chunk codes than terminal items, preserving ordinal structure (order of chunks) might be easier at this level. Second, if chunks larger than three items are formed, then not only are there fewer total chunks, but there would also be more direct links between the target and its neighbors than would be the case with chaining (with chaining two direct links are formed).

The protocol data from Experiment 2 are quite consistent with Assumption 2. Simple chunking, imaginal grouping, and story strategies, all of which induce localized rehearsal and chunking processes, were superior to displaced rehearsal strategies. Further, counting strategies wherein an item is linked with a number that represents its position in the series, was inferior to the chunking strategies, and no better than displaced rehearsal patterns. In the counting strategy only one temporally relevant association is formed.
(with the number itself). If this number code is forgotten (and the data indicate that forgetting was extensive), then because the item is not linked with temporal neighbors, there is no temporally relevant information. Hence, the counting and displaced rehearsal strategies both interfered with localized relational coding, and consequently they resulted in the two lowest mean scores in the serial-position judgment task.

This pattern of results was replicated under controlled conditions in Experiment 3. In that experiment, story strategies were superior to both, displaced rehearsal strategies, and to highly item-specific strategies (pleasantness rating group). Again, strategies that interfered with localized rehearsal patterns, resulted in poorer performance in the serial-position judgment task, than a strategy that induces localized rehearsal and chunking. The data from these experiments, therefore, do provide strong support for the central assumption of the current model.

The fact that encoding strategies had a direct influence on temporal-order learning strongly argues against the claim that this type of learning is an automatic process (Hasher & Zacks, 1979; Tzeng & Cotton, 1980). On that account the relevant processing occurs independently of attentionally-controlled strategic factors. Chosen strategies should neither enhance, or interfere with order learning.

Finally, the failure to obtain intentionality effects (subjects expecting a temporal judgment task do no better than subjects not expecting such a test), which had been considered strong support for
the automaticity hypothesis, needs to be reconsidered, in light of these findings. While in a Experiment 1 a performance advantage was obtained for the subjects trained with free-recall over subjects trained with serial-position judgments, this effect was not replicated in Experiment 2, and therefore it must be concluded that in general no overall performance differences were obtained as a function of training conditions.

The protocol analysis collected in Experiment 2 revealed that the likelihood of choosing one of the three best strategies on Trial 2 (chunking, imagery, or stories) did not differ for those subjects trained with a free-recall task, and those trained with the same serial-position judgment tasks. Thus, a large percentage of subjects in the incidental group did choose efficacious strategies, and this could account for the failure to obtain intentionality effects. Further, while subjects in the free-recall condition were more likely to employ displaced rehearsal strategies than subjects trained with a serial-position judgment task, any negative influence that this might have had on the mean performance of the former group, would be cancelled out by the fact that a similar percentage of subjects in the serial-position trained group employed equally inefficacious counting strategies. Given this pattern of choice, it is hardly surprising that the overall performance of the two groups did not differ significantly in Experiment 2.

Given the strong influence of chosen strategy, and the fact that the failure to obtain intentionality effects can be accounted for
without postulating a non-strategic account, the current findings argue, quite convincingly, that temporal-order learning should not be accorded special status in the information processing system (that is, it should not be considered an automatic process). The current model, with its emphasis on strategically-controlled, localized rehearsal patterns seems to provide an adequate account of temporal-order learning, in general.

Assumption 3.

Assumption 3 proposed that temporal judgments are made by trying to determine an item's (or pair of items) position relative to the beginning or end of the list by implicitly retrieving intervening items, or chunk codes, and counting these. This assumption is simply the logical consequence of the assumption that an item's position in a series is maintained through relational coding with other items, and not by anything like a true time tag. The current experiments do not directly test this assumption, but to the extent that the findings are consistent with the encoding assumptions of the current model, and inconsistent with the time tag notion, then a decision mechanism such as this would seem to be necessarily implied. Further, certain empirical evidence, consistent with this idea, was described above (Underwood, 1977; Tzeng, Lee & Wetzel, 1979). At present, it would seem to be reasonable to include this assumption in a model of temporal judgment performance.
Assumption 4

Assumption 4, on the other hand, was not supported by the findings from Experiments 4 and 5. This assumption proposed that when direct information regarding an item's temporal position is unavailable (due either to forgetting, or initial failure to encode), subjects will access the memory-trace strength and convert this into an estimate of how far back in the list the item occurred. Further, memory-trace strength might serve as a response bias, even when more direct information is available. A pilot study manipulated the number of category exemplars in the study-list and found no influence on paired-recency performance. According to a trace-strength account the items from categories with more exemplars would have a greater strength value than items from smaller categories, and should therefore seem more recent than the latter items. This expectation was not confirmed in the data.

Experiment 4 manipulated word frequency in a paired-recency judgment. The expectation was that low frequency items, which are more easily recognized than high frequency items, would seem to have occurred more recently than the latter items. Again, the data was inconsistent with this expectation.

Finally, Experiment 5 employed a priming manipulation. Items primed in the test stage (by using strong associates as foils), would presumambly have greater memory-trace strength than unprimed items. Therefore, on the strength account they should seem to have occurred more recently than they actually did. However, in a serial-position
judgment task, primed items were no more likely to be displaced forward (judged to have occurred later in the list than they actually did occur) than were unprimed items. Hence, no evidence of strength bias was obtained in this study.

Recall that Wells (1974) found that items that were recognized with greater confidence, were no more likely to be displaced forward than items for which confidence judgments were lower. On the strength hypothesis, items more confidently recognized would presumably have higher memory-trace strength values. The failure to find any influence of confidence on absolute recency judgment performance is inconsistent with the strength bias idea.

Taken together, all of these experiments converge on the conclusion that subjects do not appear to use memory-trace strength when making temporal judgments. Unless contrary data is subsequently obtained, it would seem to be appropriate to drop Assumption 4 from the current model. It must be concluded therefore, that if direct information regarding an item's position is unavailable, subjects will resort to guessing in a temporal judgment task.

Summary

In recent years there has been a growing interest in the question of how individuals represent temporal-order information. Various special encoding mechanisms, such as time tags have been proposed. Along with these proposals, a number of researchers have argued that temporal information is automatically encoded. Ideas of this sort
have led some theorists to postulate special temporal coding mechanisms in working memory (Shiffrin & Gillund, 1984; Detweiler & Schneider, 1988). In that way, temporal coding is accorded special status in the information processing system.

The current model proposes that temporal information is contained in the associative links between an item and its temporal neighbors; in other words, temporal position is encoded as a consequence of serial learning processes. A localized rehearsal pattern is the necessary condition in this type of learning, and rehearsal patterns are viewed as being strategically controlled. The data are quite consistent with this account, and would seem to argue against, both the time tag notion, and the automaticity hypothesis. While expectancy of an order test is not a necessary condition to induce such localized rehearsal patterns, they are nonetheless very much strategically controlled. Any further research in the area of temporal coding, must include a consideration of the kinds of rehearsal strategies that are encouraged by the specific manipulations used.

Archer, J. E., A re-evaluation of the meaningfulness of all possible CVC trigrams. *Psychological Monographs*, 1960, 74, No. 10 (Whole No. 497).


Learning and Verbal Behavior, 1973, 12, 229-238.


Moyer, R.S., & Bayer, R.H. Mental comparison and the symbolic distance effect. *Cognitive Psychology*, 1976, 8, 228-246.


APPENDIX A

Mean Lag Values for Test Pairs from Experiment 4

<table>
<thead>
<tr>
<th>Type of Test Pair</th>
<th>Mean Lag Between Items</th>
<th>Mean lag of More Recent Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>high-high</td>
<td>4.0</td>
<td>48</td>
</tr>
<tr>
<td>low-low</td>
<td>4.1</td>
<td>48</td>
</tr>
<tr>
<td>low-high</td>
<td>4.5</td>
<td>43</td>
</tr>
<tr>
<td>high-low</td>
<td>4.3</td>
<td>43</td>
</tr>
</tbody>
</table>
## APPENDIX B

**WORD AND CVC LISTS FROM EXPERIMENT 1**

<table>
<thead>
<tr>
<th>List 1</th>
<th>List 2</th>
<th>List 3</th>
<th>List 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOG</td>
<td>HAT</td>
<td>DOH</td>
<td>CEP</td>
</tr>
<tr>
<td>SEA</td>
<td>FIG</td>
<td>GIF</td>
<td>POT</td>
</tr>
<tr>
<td>SIR</td>
<td>BOX</td>
<td>TAQ</td>
<td>SUF</td>
</tr>
<tr>
<td>GAS</td>
<td>DOG</td>
<td>DUQ</td>
<td>HEJ</td>
</tr>
<tr>
<td>TIE</td>
<td>GUM</td>
<td>JOP</td>
<td>HAK</td>
</tr>
<tr>
<td>JAR</td>
<td>EAR</td>
<td>VOD</td>
<td>VID</td>
</tr>
<tr>
<td>FAN</td>
<td>FLY</td>
<td>FOW</td>
<td>BEM</td>
</tr>
<tr>
<td>DAD</td>
<td>OAK</td>
<td>KUL</td>
<td>JIV</td>
</tr>
<tr>
<td>TIP</td>
<td>PIT</td>
<td>ROH</td>
<td>QIN</td>
</tr>
<tr>
<td>COP</td>
<td>NUT</td>
<td>NEQ</td>
<td>MIP</td>
</tr>
<tr>
<td>LID</td>
<td>BUS</td>
<td>DAK</td>
<td>YUC</td>
</tr>
<tr>
<td>BOW</td>
<td>CAP</td>
<td>QAD</td>
<td>QAL</td>
</tr>
<tr>
<td>PIP</td>
<td>ICE</td>
<td>HIV</td>
<td>KAC</td>
</tr>
<tr>
<td>LIP</td>
<td>MUD</td>
<td>FID</td>
<td>FAM</td>
</tr>
<tr>
<td>JAW</td>
<td>JET</td>
<td>NAZ</td>
<td>PEZ</td>
</tr>
<tr>
<td>ROD</td>
<td>ERA</td>
<td>SAH</td>
<td>SOH</td>
</tr>
<tr>
<td>QIL</td>
<td>Leg</td>
<td>MAB</td>
<td>LUF</td>
</tr>
<tr>
<td>HAM</td>
<td>SUM</td>
<td>MUZ</td>
<td>MOK</td>
</tr>
<tr>
<td>FEE</td>
<td>COW</td>
<td>NES</td>
<td>GUK</td>
</tr>
<tr>
<td>HEN</td>
<td>PIE</td>
<td>KIZ</td>
<td>PAJ</td>
</tr>
<tr>
<td>CAT</td>
<td>GUY</td>
<td>BIV</td>
<td>LEB</td>
</tr>
<tr>
<td>PEN</td>
<td>VAN</td>
<td>MUZ</td>
<td>NEF</td>
</tr>
<tr>
<td>GIN</td>
<td>CUP</td>
<td>PUH</td>
<td>PUQ</td>
</tr>
<tr>
<td>BAT</td>
<td>NET</td>
<td>LOH</td>
<td>LAH</td>
</tr>
<tr>
<td>LAP</td>
<td>BAG</td>
<td>SOQ</td>
<td>CIF</td>
</tr>
<tr>
<td>BAR</td>
<td>BAY</td>
<td>NUS</td>
<td>QIP</td>
</tr>
<tr>
<td>ARM</td>
<td>KID</td>
<td>YOM</td>
<td>VIZ</td>
</tr>
<tr>
<td>PAN</td>
<td>POT</td>
<td>BAQ</td>
<td>WOH</td>
</tr>
<tr>
<td>KEY</td>
<td>SKY</td>
<td>HEF</td>
<td>HEB</td>
</tr>
<tr>
<td>PIN</td>
<td>TEA</td>
<td>QIK</td>
<td>POZ</td>
</tr>
</tbody>
</table>
# APPENDIX C

**TEST-PAIRS FROM EXPERIMENT 4**

<table>
<thead>
<tr>
<th>High-High</th>
<th>Low-Low</th>
<th>Low-High</th>
<th>High-Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAR-MONEY</td>
<td>BRIAR-RINK</td>
<td>HOTHOUSE-WATER</td>
<td>BODY-LEASH</td>
</tr>
<tr>
<td>CHURCH-BLOOD</td>
<td>HALLWAY-TACK</td>
<td>SULTAN-PICTURE</td>
<td>SQUARE-LATHE</td>
</tr>
<tr>
<td>HOUSE-FEET</td>
<td>HOOF-TAPESTRY</td>
<td>BUGLE-OFFICE</td>
<td>HEAD-HEXAGON</td>
</tr>
<tr>
<td>DOOR-PAPER</td>
<td>ICICLE-TOAD</td>
<td>ORPHANAGE-LETTERS</td>
<td>COLLEGE-JAWBONE</td>
</tr>
<tr>
<td>SCHOOL-WOMAN</td>
<td>LENTILS-LICE</td>
<td>HELMET-BOOK</td>
<td>PRESIDENT-FOIL</td>
</tr>
<tr>
<td>FIRE-COURT</td>
<td>MUSTARD-STRAPS</td>
<td>HYENA-STATE</td>
<td>BUSINESS-CRATER</td>
</tr>
<tr>
<td>GUN-NATURE</td>
<td>FLASK-NEBULA</td>
<td>QUARRY-STREET</td>
<td>CHILDREN-DIGIT</td>
</tr>
<tr>
<td>HEART-EARTH</td>
<td>MURAL-NIGHTSHIRT</td>
<td>BATON-HUSBAND</td>
<td>BED-LAVA</td>
</tr>
</tbody>
</table>

149
## APPENDIX D

### ASSOCIATES USED IN EXPERIMENT 5

<table>
<thead>
<tr>
<th>Target</th>
<th>Prime</th>
<th>Target</th>
<th>Prime</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOWN</td>
<td>CITY</td>
<td>BUTTER</td>
<td>BREAD</td>
</tr>
<tr>
<td>QUEEN</td>
<td>KING</td>
<td>SHORT</td>
<td>LONG</td>
</tr>
<tr>
<td>TRUCK</td>
<td>CAR</td>
<td>SKIP</td>
<td>JUMP</td>
</tr>
<tr>
<td>NOW</td>
<td>THEN</td>
<td>NEEDLE</td>
<td>THREAD</td>
</tr>
<tr>
<td>DIE</td>
<td>LIVE</td>
<td>TUB</td>
<td>BATH</td>
</tr>
<tr>
<td>SUN</td>
<td>MOON</td>
<td>SHALLOW</td>
<td>DEEP</td>
</tr>
<tr>
<td>SWEET</td>
<td>SOUR</td>
<td>WOOL</td>
<td>SHEEP</td>
</tr>
<tr>
<td>WEB</td>
<td>SPIDER</td>
<td>DREAM</td>
<td>SLEEP</td>
</tr>
<tr>
<td>GRASS</td>
<td>GREEN</td>
<td>LOSE</td>
<td>FIND</td>
</tr>
<tr>
<td>GIRL</td>
<td>BOY</td>
<td>CATS</td>
<td>DOGS</td>
</tr>
<tr>
<td>LOW</td>
<td>HIGH</td>
<td>NAIL</td>
<td>NAMMER</td>
</tr>
<tr>
<td>HARD</td>
<td>SOFT</td>
<td>PEPPER</td>
<td>SALT</td>
</tr>
<tr>
<td>WINDOWS</td>
<td>DOORS</td>
<td>NURSE</td>
<td>DOCTOR</td>
</tr>
<tr>
<td>TIGER</td>
<td>LION</td>
<td>SHOE</td>
<td>FOOT</td>
</tr>
<tr>
<td>DARK</td>
<td>LIGHT</td>
<td>ROUGH</td>
<td>SMOOTH</td>
</tr>
<tr>
<td>RUG</td>
<td>CARPET</td>
<td>CHAIR</td>
<td>TABLE</td>
</tr>
</tbody>
</table>