

INFORMATION TO USERS

This manuscript has been reproduced 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.

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

**A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600**

**CLOSING IN ON THE ROLE OF ATTENTION
IN IMPLICIT SEQUENCE LEARNING**

by

Andrew T. Hsiao

**A dissertation submitted to the Graduate Faculty in Psychology in
partial fulfillment of the requirements for the degree of Doctor of
Philosophy, The City University of New York**

1997

UMI Number: 9807940

**Copyright 1997 by
Hsiao, Andrew T.**

All rights reserved.

**UMI Microform 9807940
Copyright 1997, by UMI Company. All rights reserved.**

**This microform edition is protected against unauthorized
copying under Title 17, United States Code.**

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

© 1997

ANDREW T. HSIAO

All Rights Reserved

This manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

September, 5, 1997 _____
Date Chair of Examining Committee

9-9-97 _____
Date Executive Officer

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract**CLOSING IN ON THE ROLE OF ATTENTION
IN IMPLICIT SEQUENCE LEARNING**

by

Andrew T. Hsiao

Advisor: Professor Arthur S. Reber

Does implicit learning require attention? This is the question the present study attempted to address. Nissen and Bullemer (1987) first raised the question by testing subjects in a serial reaction time (SRT) task performed together with a tone-counting secondary task. In one group, the SRT task followed a repeating pattern while in the other, it was random. The secondary task was introduced to engage the attentional mechanism, thus forcing the SRT task to be performed when less attention was available. As the difference in RT between the repeating-pattern group and the random group failed to reach statistical significance, they concluded that sequence learning requires attention. However, the conclusion has been challenged in subsequent studies using different approaches (Cohen, Ivry & Keele, 1990; Curran & Keele, 1993, Frensch, Buchner & Lin, 1994).

The present study showed that, in contrast to Cohen et al. (1990), even sequences made up entirely of ambiguous transitions can still be learned. However, instead of coming to a conclusion that sequence learning may be

achieved without attention, further analyses indicated that things are far more complicated than researchers in the field initially realized. First, it was found that subjects who performed the secondary task relatively poorly showed much better learning than those who did well. If doing well on the secondary task implies paying more attention to it, this would suggest that sequence learning is affected by the availability of attention. The fact that giving more attention to the secondary task resulted in a compromise in sequence learning indicates that (1) the two tasks share common resources; and (2) these resources are limited in capacity. Second, it was found that when subjects were given a practice on the secondary task prior to performing the dual-task SRT task, sequence learning was reduced. It is likely that the practice emphasized the difficulty of the secondary task. This encouraged the subjects to allocate more attention to the task, thus making attention less available to sequence encoding. Third, manipulating the onset of the secondary task stimulus systematically demonstrated that learning was impaired when subjects were paying more attention to the secondary task while the two competing tasks were presented in close succession. This suggests that the two cognitive processes can share the same resources successfully unless both tasks demand attention at the same time or in a very narrow time frame. Together, these findings invite the conclusion that attention remains a critical determinant in sequence learning.

ACKNOWLEDGEMENT

In the standard professional vita, earning one's Ph.D. occupies but one line, which simply specifies the year in which the degree was conferred. I wish the experience is as simple as that. But for many, particularly myself, that is an understatement in cosmological terms. While I was celebrating the end of this journey, I took some time to relive the experience mentally. Weighing the gain and loss due to the efforts made to accomplish this step, I might say the benefit still tips the scale. No doubt, the goal was achieved with the help and support of many individuals who, in his or her own way, made a difference in shaping my experience.

First and foremost, I am deeply in debt to my mentor, Prof. Arthur Reber. Anyone who deigns to read this treatise should not fail to recognize his theoretical influence. Over the years, not only did he teach me how to do good research, he also emphasized the importance of "seeing the forest" from the little niche of research I was confined in. In addition, he has always impressed me by his willingness to help his students, even if this meant going out of his way.

I am also grateful to Prof. Andy Delamater for the many hours he generously spent with me in discussing methodological as well as theoretical issues. Had not for his wise guidance at critical times, I might have entered several blind alleys before treading on the right path for the project.

Thanks are also due to Profs Don Scarborough, Benzion Chanowitz and Kevin Sailor. Their critical comments and suggestions helped weeded many weaknesses and errors in earlier versions of this dissertation.

The list of creditors will not be complete without mentioning the names of a number of scholars who have contributed significantly to our understanding of implicit sequence learning. These include Profs Steve Keele, Mike Stadler, Tim Curran, Asher Cohen, Peter Frensch, and Dan Willingham. I would also like to mentioned the names of two other scholars who worked outside this immediate circle. They are Profs Hal Pashler and David Shanks. Thanks to the convenience in communication via e-mail, I was able to benefit from these individuals of their experience and expertise on relevant issues literally on a global level.

Of course, errors of various kinds likely remain even in this final version of the dissertation. However, I am solely responsible for the lack of foresight in cases when some statements prove to be incorrect in future, or for overlooking obvious mistakes which could be rid of had more care been exercised.

Lastly, two sources of spiritual support sustained me to bring the long journey to the completion even though the contents of the dissertation would not be meaningful to one or would only make sense to the other years later. I am referring to my mother, who lost a bet that she would outlive my project, and my daughter, who has grown up with it.

TABLE OF CONTENTS

CHAPTER 1. THEORIES AND MECHANISMS OF SEQUENCE LEARNING UNDER DISTRACTION	1
1.0 Introduction	1
1.1 Sequence Characteristics And Attentional Demand	3
1.1.1 Sequence length	3
1.1.2 Uniqueness and ambiguity in event transitions	5
1.1.3 Overall statistical properties	7
1.1.4 Higher-order contingencies	8
1.2 The Role Of The Secondary Task Plays In Sequence Learning	12
1.2.1 The secondary task handicaps the hierarchical encoding mechanism	12
1.2.2 The secondary task suppresses expression of learning	14
1.2.3 The secondary task disrupts grouping in sequence learning	16
1.2.4 The secondary task reduces short-term memory capacity	20
1.2.5 The secondary task adds complexity to the learning task	22
 CHAPTER 2. METHODOLOGICAL CONSIDERATIONS	 27
2.1 Demonstrating Learning	27
2.2 Demonstrating Awareness	32
2.3 Manipulating Available Attention	37
 INTERLUDE: AN ONGOING SUMMARY AND SOME QUESTIONS	 47
 CHAPTER 3. EXPERIMENT 1, EFFECT OF RESIDUAL ATTENTION AS A SUBJECT FACTOR	 52
3.1 Experiment 1a	52
3.1.1 Methods	53
3.1.2 Data screening	63
3.1.3 Documenting learning	63
3.1.4 Results and discussion	68
3.2 Experiment 1b	84
3.2.1 Methods	84
3.2.2 Data screening	85
3.2.3 Results and discussion	86
3.3 Chapter Summary And Conclusions	97
 CHAPTER 4. EXPERIMENT 2, EFFECT OF ATTENTION AS A FUNCTION OF TASK DEMAND	 101

4.1 Experiment 2a	102
4.1.1 Methods	102
4.1.2 Data screening	104
4.1.3 Results and discussion	106
4.2 Experiment 2b	110
4.2.1 Methods	112
4.2.2 Data screening	112
4.2.3 Results and discussion	113
4.3 Chapter Summary And Conclusions	115
CHAPTER 5. EXPERIMENT 3, SEQUENCE LEARNING AND ATTENTION SHARING	124
5.1 Methods	127
5.2 Data Screening	130
5.3 Results And Discussion	130
5.4 Chapter Summary And Conclusions	141
CHAPTER 6. GENERAL DISCUSSION AND CONCLUSIONS	143
6.1 A Summary Of The Findings From This Study	143
6.2 Some Theoretical And Empirical Issues Of Attention	145
6.3 The Relationship Between Attention And Awareness	148
REFERENCES	150

LIST OF TABLES

Table 3.1 Sequences used in Exp. 1a and their statistical properties (simple event frequencies and 1st-order conditionals).	57
Table 3.2 Design features Experiment 1a.	59
Table 3.3 The grammatical sequence used in Exp. 1b and its statistical properties (simple event frequencies and 1st-order conditionals). The transfer sequence is the same as used in Exp. 1a (see Table 3.1).	86
Table 3.4 Design features Experiment 1b.	87
Table 4.1 The statistical properties (simple event frequencies and 1st-order conditionals) of the (Rep and Gram) sequences used in Exps 2a and 2b.	103
Table 5.1 Training and transfer sequences used in Experiment 3 and two properties of these sequences: simple event frequencies and 1st-order conditionals.	128

LIST OF CHARTS AND DIAGRAMS

- Figure 2.1 Stimulus Onset Asynchrony (SOA) in a dual-task setting: temporal relationship between the onset of the primary task and that of the secondary task.** 39
- Figure 3.1 a. A diagram of the stimulus features; b. Layout of response keys on the numerical key pad.** 54
- Figure 3.2 An illustration of the temporal relationship between onsets of the secondary stimulus and the primary stimulus. The boxes indicate psychological processes invoked by the secondary stimulus.** 56
- Figure 3.3 Distribution of average error rate on the secondary target-signal counting task (E_{R2}) in Experiment 1a before data were screened.** 66
- Figure 3.4 RTs over the trial blocks are plotted by how subjects attended to the secondary task and the modality of the secondary stimulus.** 70
- Figure 3.5 Direct (explicit, $Pred$) and indirect (implicit, $RT_7 - RT_6$) learning measures are plotted to demonstrate the relation between the two measure.** 75
- Figure 3.6 RTs associated with each of the stimulus components are plotted across the trial blocks by sequence type. Panel a for Rep sequence and Panel b for Gram sequence.** 76
- Figure 3.7 Size of learning effect associated with each of the stimulus components are shown by Good Counters (Left) and Poor Counters (Right).** 79
- Figure 3.8 RTs plotted for first-order conditionals appeared in both the training and transfer sequence, measured in Block 6 (Old6) and Block 7 (Old) and those only appeared in the transfer sequence (New).** 80
- Figure 3.9 A diagram of the artificial grammar used to generate a sequence of elements to determine the order in which the SRT task stimulus is presented.** 85
- Figure 3.10 Distribution of average error rate (E_{R2}) on the secondary target-signal counting task in Exp 1b before screening was applied.** 88

- Figure 3.11 RTs over the trial blocks are plotted by how subjects attended to the secondary task and the modality of the secondary stimulus. 88
- Figure 3.12 Direct (explicit, $Pred$) and indirect (implicit, $RT_7 - RT_6$) learning measures are plotted to demonstrate the relation between the two measure. 91
- Figure 3.13 RTs associated with each of the stimulus components are plotted across trial blocks. 93
- Figure 3.14 Learning effects associated with each of the stimulus components are shown by Good Counters and Poor Counters. 94
- Figure 3.15 RTs plotted for first-order conditionals appeared in both the training and transfer sequence, measured in Block 6 (Old6) and Block 7 (Old) and those only appeared in the transfer sequence (New). 96
- Figure 4.1 RTs are plotted by secondary stimulus modality and by how the secondary task was performed. Left panel: with a repeating pattern; Right panel: with a grammatical sequence. 107
- Figure 4.2 RTs are plotted by secondary stimulus modality and by how the secondary task was performed. Left panel: with a repeating pattern; Right panel: with a grammatical sequence. 114
- Figure 4.3 Plotted are mean secondary task error rates (E_{R2}) across the trial blocks in Experiments 2a (filled squares) and 2b (empty squares). 116
- Figure 4.4 a. Distribution of averaged secondary task error rate in Experiment 2a; b. Distribution of averaged secondary task error rate in Experiment 2b. 117
- Figure 4.5 RTs for Good Counters and Poor Counters are plotted across the trial blocks for data from Exp 2a (without secondary task practice) and Exp 2b (with practice). 118
- Figure 4.6 Effect of learning as a function of availability of attention controlled by internal and external conditions. 119
- Figure 4.7 Plotted are RTs by type of trio structure defined by whether it contains a reversal. 120

- Figure 5.1** RTs are plotted across the trial blocks in a set of three graphs by Good Counters, Poor Counters and Estimators respectively. 132
- Figure 5.2** Sequence learning effects are plotted by allocation of attention and the onset of the secondary stimulus. 133
- Figure 5.3** Figure reproduced based on data shown in Fig 5.1 with the data in Blocks 7 and 8 removed. The original Blocks 9 and 10 are now labeled 7 and 8. 135
- Figure 5.4** The implicit learning scores ($RT_9 - RT_8$) and the two explicit knowledge scores (Pred and Prod) are plotted in three graphs to show the relationship between each pair. 138
- Figure 5.5** RTs are plotted across the trial blocks according to whether subjects were aware of the presence of a regular pattern. 139
- Figure 5.6** Relationship between sequence learning effect ($RT_7 - RT_8$) and confidence level (High vs. Low) for secondary task response. 140

CHAPTER 1
THEORIES AND MECHANISMS OF SEQUENCE LEARNING
UNDER DISTRACTION

1.0 Introduction

The present dissertation attempts to address a simple question: Does implicit sequence learning require attention? To avoid getting entangled hopelessly in philosophical pitfalls, the concepts here are all empirically defined. In this context, sequence learning refers to a process by which subjects acquire information about the structural constraints of a sequence of stimulus events in a serial reaction time (SRT) task. Sequence learning is said to be implicit if (1) the particular learning process is not attempted intentionally by the subject, and/or (2) the product of such learning remains beyond conscious recollection. Unintentional learning is achieved procedurally by presenting the stimulus in an incidental learning situation. Nonconscious knowledge is checked against verbal protocol and recall/recognition measures. Learning is indexed by performance changes that are consistent with the presence or absence of sequence constraints.

All of the studies we will discuss here use one or another variation on a standard SRT procedure first developed by Nissen and Bullemer (1987). In this procedure subjects are instructed to press as rapidly as possible one of four keys (or on

occasion, more) in response to a corresponding simple visual stimulus target (such as a dot or an asterisk) as it appears at one of four (or more) locations arranged horizontally across the computer screen. The location of the stimulus is determined either by a sequence which repeats over and over (e.g., DBCACBDCBA is the one used by Nissen & Bullemer, 1987) or one that is nonrepeating but follows a set of rules (see, e.g., Cleeremans & McClelland, 1991). Sequence learning is assessed either by comparing RTs to the training sequence with those obtained on a nonstructured or random sequence, or by introducing a block of trials late in training where the sequence is replaced with a random or novel sequence. In the former situation where a between-subjects design is used, subjects who received a structured sequence are expected to respond faster than subjects who received a random sequence. In the latter with a within-subjects design, RTs should increase when the sequence is changed. The role of attention is typically explored by introducing a secondary task, most commonly a tone-counting task in which following each target stimulus either a high or low-pitched tone sounds and subjects are required to keep a running count of one of them.

Using this procedure, Nissen and Bullemer (1987) first raised the question of whether implicit sequence learning was demanding of attentional resources. Nissen and Bullemer's research was based on the assumption of a limited central capacity in human information processing. In a series of experiments, they found that when performing the SRT task alone, subjects who responded to the repeating pattern displayed reliably faster RTs than those who responded to a random sequence. However, when a secondary tone-counting task was introduced, RTs for the two

conditions did not differ significantly prompting Nissen and Bullemer to conclude that sufficient attention is a critical condition for implicit sequence learning. In a subsequent experiment, Korsakoff amnesics, in the absence of the distractor, also succeeded in learning the sequence but were apparently unaware of what was learned, suggesting that conscious memory is not necessary for such learning.

The dissociation between attention and conscious awareness of a learning experience attracted a good deal of attention for at least two reasons. First, it made clear that a satisfactory account of attention will have to extend beyond topics concerned primarily with the apprehension of external stimulation such as visual attention, focused or selective attention, and divided attention (Keele, 1973; Treisman, 1969; Treisman & Gelade, 1980) and take into account the role of attention in implicit, unconscious cognitive processes. Second, it cast new light on previous findings that automatic processes such as logging the frequency of events in stimulus displays required few if any attentional resources (Hasher & Zacks, 1984). In this chapter we review the now extensive literature on the role of attention in implicit sequence learning and examine a number of theories developed to account for the findings.

1.1 Sequence Characteristics And Attentional Demand

An early development in the research of the role of attention in implicit sequence learning was the finding that various characteristics such as the length and statistical structure of the sequence place different demands on attention. Several of these studies will be discussed in some detail in this section.

1.1.1 Sequence length

One sequence characteristic is its sheer length. Although different researchers have used sequence lengths varying from 5 elements (Cohen, et al., 1990) to the length of an entire block of 96 trials (Reed & Johnson, 1994), Howard and her colleagues (Howard & Howard, 1989; Howard, Mutter, and Howard, 1992) were the first to examine the specific effect of sequence length on learning. They trained their subjects with either a 10-element (DBCACBDCBA) sequence or a 16-element (BCACDCDBADADACAD) sequence. The results revealed a more pronounced learning effect associated with the shorter sequence. Howard and Howard hypothesized that the longer sequence required more attentional capacity to encode because it had to be analyzed into more chunks to resolve inherent ambiguities. Similar findings were obtained by Pascual-Leone, Grafman, Clark, Stewart, Massaquoi, Lou, and Hallet (1993) although they argued that longer sequences, simply by virtue of having more elements, place a greater demand on working memory. However, since other statistical properties were left uncontrolled (most notably, the basic frequency of occurrence of each component and the number of unique transitions both of which are confounded with the effect of sequence length), both conclusions should be viewed with caution.

More recently, Stadler and Neely (in press), realizing that it is difficult to control for all the relevant characteristics of a sequence at the same time, attempted to unpack the length vs. structure question by using different combinations of sequence length and levels of structural redundancies. In a series of three experiments, they found the latter to be the better predictor of implicit sequence learning. Specifically, a longer sequence with "higher structure" can be better learned than a shorter one with "lower

structure" where structure can be taken to represent a metric akin to the classic, Shannon and Weaver notion of information. Gauged by this information yardstick, highly structured sequences are ones whose subsequent events are highly predictable given the immediately preceding event(s). As there will more discussions on the notion of structure later, suffice it to say here that sequence length alone cannot be viewed as a significant determinant of sequence learning.

1.1.2 Uniqueness and ambiguity in event transitions

Cohen, Ivry, and Keele (1990) investigated implicit sequence learning with three types of sequences having different inner structures, which they referred to as *ambiguous*, *unique* and *hybrid* respectively. An ambiguous sequence is one in which the occurrence of one event provides insufficient information about the succeeding event (e.g., in ACDBCABD, knowing A has occurred, one cannot be sure whether it will be followed by B or C). In contrast, if a sequence contains events each of which is determined reliably by the event immediately preceding it, it is called a unique sequence (e.g., in ADCBADCB, A is always followed by D, which, in turn, is always followed by C, and so on). A hybrid sequence, as the term suggest, is simply one that is made up of a mixture of unique and ambiguous transitions. In contrast to Nissen and Bullemer's (1987) conclusion, Cohen et al. found that sequence learning under secondary task distraction is impaired only when the sequence is of the ambiguous type--which happened to be the sequence type used by Nissen and Bullemer.

Cohen et al. (1990) theorized that unique sequences, which can be encoded as simple pair-wise associations, may be learned automatically, whereas ambiguous

sequences, which call for hierarchical encoding, require attention. This led them to propose a two-mechanism theory in which a non-attentional mechanism is sufficient for associative learning whereas an attentional mechanism must be invoked for encoding information hierarchically. This theory was buttressed by Curran and Keele (1993) who demonstrated that the non-attentional form of learning was operational in the presence or absence of the tone-counting task while the attentional form appeared to operate only when freed from distraction. However, more recent work has complicated matters. Frensch, Buchner, and Lin (1994) reported that subjects were able to learn ambiguous transitions within the context of a hybrid sequence even in a dual task situation. Hsiao (1995) observed sequence learning in a dual-task setting for an ambiguous sequence, including a nonrepeating sequence generated by an artificial grammar along the lines of the one used by Cleeremans and McClelland (1991). Other replications also yielded results that differed from the original findings (Keele, personal communication, 1995; Keele & Jennings, 1992).

It is not entirely clear why the picture changed as research on sequence learning progressed. One possibility is that measures of learning have attained greater sensitivity due to improved control—for example, by using pseudo-random transfer sequences that matched with the practice sequences in simple event frequencies. Another may be that subtle differences in instructions given to participants prior to practice on the sequence are playing a role (Hsiao, 1997). Yet another are the different data inclusion/exclusion criteria adopted by different researchers. For example, subjects with systematically long RTs (outliers) are usually discarded, but the criterion for dropping them is not always the same. In addition, only the data

from subjects whose accuracy on the tone-counting task exceeded certain arbitrary percentage are used, however, not everyone uses the same cut-point. In spite of these procedural and methodological complications, the evidence is sufficiently strong to conclude that, other factors notwithstanding, even sequences made up entirely of non-unique transitions may be learned under conditions of limited attention although they are surely learned more slowly than unique sequences.

1.1.3 Overall statistical properties

Stadler (1992), arguing that learning likely extends beyond the detection of simple event-to-event covariations, emphasized the importance of overall statistical structure. He varied the structural properties of learning sequences by creating three sequences, labeled as having *low*, *medium* or *high* statistical structure. Specifically, he joined a 6-element root string (e.g. BDBCAB) with one of three suffix strings (e.g. ADAC, CDBC, or DBCD), to build three sequences, BDBCABADAC, BDBCABCDBC, and BDBCABDBCD, which correspond to the low, medium and high structures respectively. Note that the addition of the suffix in the low structure case only created new pair-wise associations; in the medium structure case it produced three instances of existing pair-wise associations (two BCs and a DB) and an instance of trio-repetition (DBC) that already existed; in the high structure case four instances of pair- and two instances of trio-repetitions were created (BD, two DBs, BC; BDB, DBC).

It was posited that, with the increase of structural constraints, certain event sequences become more predictable by virtue of the fact that these events have a greater probability of occurrence after certain other events. Indeed, the results showed differences in sequence learning commensurate with the three different levels of

manipulation. The effect on learning is similar to what was found in Cohen et al. (1990) in that a sequence that contains highly probable event transitions is learned more easily. If the three sequences are analyzed in Cohen et al.'s terms, it can be seen that the low-structure sequence is composed of 10 transition types, none of which is unique; the medium-structure sequence is made of 7 transition types with two unique transitions; the high-structure sequence contains 6 transition types with two unique transitions, one of which is repeated 3 times. If we take the structural analysis a step further to consider transitions of events when two preceding events are given, which forms event-trios (e.g., BDC), or second-order conditionals, we find 10 types of such higher-order transitions in the low- and medium-structure sequences but only 8 types in the high-structure sequence.

However, Stadler only used the single-task paradigm. It remains to be seen how subjects would perform with such sequences when attention is diverted. Nevertheless, the primary conclusion seems secure. The fewer the constraints the more attentional resources will be required to learn that sequence.

1.1.4 Higher-order contingencies

Cohen et al. (1990) had argued that ambiguous sequences are difficult to learn because higher-order contingencies must be encoded in order to resolve the ambiguity. Given time and sufficient attention, people can, and usually do, learn higher-order contingencies of events which enable them to act effectively. Cleeremans and McClelland (1991) reported that subjects were able to benefit from sequential dependencies up to the fourth order and Millward and Reber (1972), using the probability learning procedure, found sensitivities to events occurring seven trials

back. The question now is whether this kind of sensitivity can be attained when attention is diverted.

Reed and Johnson (1994) were the first to address this issue using the SRT task. Unlike in the standard procedure in which the stimulus locations are determined by an n -element sequence recycled k times in a block, they used a long sequence which was created by concatenating eight different 12-event sequences that shared the same set of second-order conditionals (SOC). In such a sequence, each event becomes predictable given two other events preceding it. After extensive training (17 blocks) half the subjects were transferred to a different sequence that had the same SOCs while the other half received a pseudo-random sequence. The pseudo-random sequence had transitions selected randomly but similar simple event frequencies as the sequence used in training. Reed and Johnson found that the sequence-transfer subjects showed significant advantage over the pseudo-random-transfer subjects implying that second-order dependencies could be learned even when attentional resources were diverted.

However, it may still be too early to consider the issue resolved. Recently, Heuer and Schmidtke (1996) conducted a study in which different groups of subjects performed a SRT task either with or without a distractor. For some subjects, the behavior of the stimulus could be predicted accurately from the preceding event (i.e., with unique transitions); for others the stimulus becomes deterministic only when two or three events before it were fixed (with higher-order contingencies). Sequence learning was assessed both in a single-task and or a dual-task setting. The results run counter to Reed and Johnson's findings; only first-order dependencies seemed to have been learned when the SRT task was practiced under distraction. Learning of the

second-order contextual information was impaired even when both practice and test were done without distraction. One reason for the failure of replication may be due to the fact that Heuer and Schmidtke's subjects did not receive as much training as those in Reed and Johnson's study. Obviously more work is still needed to inform the question, not whether higher-order event dependencies can, or cannot, be learned in the presence of a distractor task, but under what circumstances they can be learned. This issue is important because its resolution will help provide understanding of the relationship between mechanisms that handle explicit processes and those that mediate implicit processes.

To sum up, recent research on implicit sequence learning seems to show that the SRT task is sensitive to manipulations on the sequence characteristic variable. Of particular interest for our discussion in this section is that the sequence-type variable seems to interact with the availability of attention. Nissen and Bullemer's (1987) initial conclusion regarding attentional requirements with respect to sequence learning under dual task condition has proven to be premature. Cohen et al.'s (1990) study showed that Nissen and Bullemer failed to find learning because they had restricted training to the ambiguous type of sequence. In addition, had Nissen and Bullemer used the between-subjects design as did Cohen et al. and most others subsequently, they might also have found learning in their dual-task subjects. I'll have more to say about this issue in the next chapter.

Moreover, as the studies reviewed in this section demonstrate, even Cohen et al.'s discovery appears to be less than conclusive. Later research yielded evidence that even the ambiguous type of sequence can be learned under attentional distraction

using the dual-task procedure, although sequences containing unique transition were generally learned better. In their original report, Cohen et al. suggested that more complex sequential structure was best encoded using hierarchical representation whereas simpler structures can be learned via paired association. However, since later work showed that ambiguous transitions can also be learned, there seems to be no compelling reason to believe that the hierarchical encoding mechanism is completely inhibited when attention is diverted.

One difficulty in understanding the role of attention in sequence learning stems from a lack of a clear definition of attention itself. If we assume that (a) attention is a central processor with a limited capacity and (b) encoding information of various complexity requires more or less attention, we may attribute impairment in, or failure of, sequence learning to this central bottle-neck when competing tasks are performed simultaneously. But there are alternatives. It is conceivable that sequence learning that calls for more complex processing may still be possible in a dual-task setting because the secondary task may not deplete the attentional resources completely, or the task may not demand attention at all times. Thus the sequence encoding mechanism might take advantage of the residual processing capacity in a phenomenon that can be called attention-sharing and/or attention-switching (Curran & Keele, 1992; Pashler & Carrier, 1996; Stadler, 1995).

It follows from this perspective that sequence learning will be achieved in a graded rather than all-or-none fashion. However, if different tasks are handled by parallel and independent mechanisms or processing modules (Curran & Keele, 1993), and if the function of attention is to monitor or coordinate various processes, then it

may be possible to observe a dissociation between success or failure of task performance and overall system load. Progress in cognitive neuroscience in recent years is illuminating such issues by monitoring blood flow and consumption in different parts of the brain while cognitive tasks are being performed. But greater resolution is needed to track down the flow of information, or neural impulses, throughout the brain. Until then, the jury is still out on the exact role of attention in sequence learning.

1.2 The Role of The Secondary Task In Sequence Learning

As was shown in Nissen and Bullemer's (1987) original study and all subsequent work, there is little doubt that performing a secondary task concurrently is disruptive because it invariably slows down subjects' responses to the primary SRT task. But what exactly this task disrupts is under considerable debate. In this section, I'll review research that attempts to uncover the nature of the disruption by the secondary task. In what follows I will need to use the term "learning effect", or sometimes, "transfer effect", to refer to a significant increase in RT on a block of trials in which the underlying structure of the sequence has been altered in some manner.

1.2.1 The secondary task handicaps the hierarchical encoding mechanism

Cohen et al. (1990) observed that subjects were able to learn an ambiguous sequence when the SRT task was performed alone but were unable to do so under dual-task conditions. They hypothesized, as discussed in the preceding section, that sequences with unique transitions can be learned by a mechanism that operates on associative principles whereas sequences with ambiguous transitions require one capable of hierarchical encoding. Since Cohen et al. assumed that performing the

tone-counting task reduced attention available to the encoding of structured sequences, the failure to learn ambiguous sequences indicates that the hierarchical mechanism requires attention to operate.

Curran and Keele (1993) further explored this dual mechanism theory in a series of experiments. They referred to the two types of learning involving hierarchical encoding and associational encoding as the *attentional* and *non-attentional* forms respectively. In one experiment their subjects first performed the SRT task without distraction. The stimulus followed a repeating pattern but changed to a random sequence in a transfer procedure to assess learning. They repeated this procedure once more, but in a dual task setting this time. As it turned out, the learning effect was roughly twice as much in the former as in the latter. In a second experiment subjects started off under a dual-task condition. Learning was first measured in the presence of the distractor and then when it was absent. Surprisingly, the single task advantage observed in the former experiment disappeared in the latter; the amount of learning was the same whether measured in a single or a dual task setting. Curran and Keele interpreted this result as meaning that both the attentional and the non-attentional forms of learning occur when sufficient attention is available in the training phase, but only the non-attentional form operates when resources are diverted by a distraction task. Specifically, the tone-counting task reduces attentional-resources necessary for parsing and encoding sections of an ambiguous sequence. Although later research showed, as I mentioned above, that performing a secondary task may not necessarily preclude the learning of ambiguous sequences, existing data still support their basic argument that encoding more complex structural information requires more attention.

1.2.2 The secondary task suppresses expression of learning

There may be alternative explanations for the results observed by Curran and Keele (1993), however. One possibility is that the distraction task operates in a way that impairs expression of what has been learned. For instance, in their first experiment where subjects practiced the SRT task in a single task condition, learning was more pronounced when assessed in the single than in the dual task condition. In another experiment, Curran and Keele had one group practice with the distractor and another without. Here, both groups exhibited similar levels of sequence learning. It is entirely possible that sequence knowledge acquired in a single task condition failed to gain expression when the distractor task was added.

This notion has been embraced by Frensch, Lin, and Buchner (1997). In one experiment, they manipulated the amount of practice subjects received under single and dual task conditions and observed the effect of practice in the presence or absence of a distraction. All subjects received the same amount of practice on the primary SRT task. However, some subjects received more practice in a single task situation whereas others received more practice in a dual task situation. Specifically, three groups were run under three practice schedules, indexed by the number of blocks practiced in single (ST) and dual task (DT) conditions (i.e., 6DT/1ST, 4DT/3ST, & 2DT/5ST respectively) before learning was assessed. The 6DT/1ST group, for instance, designates the group that first performed 6 blocks of the SRT task with a distractor, followed by 1 block without the distractor. Learning was evaluated in a single-task situation for all three groups in this experiment. Frensch et al. found that regardless of which practice schedule was used the amount of learning was virtually

identical.

In another experiment with a slightly different procedure, all subjects began by performing the SRT task in a single task condition, and then were shifted to a dual task at different times in the training phase (i.e., the three groups were: 6ST/1DT, 4ST/3DT, & 2ST/5DT). Here response rate was proportional to the amount of practice subjects received under the dual task condition (for instance, the 2ST/5DT group had the fastest overall RTs in the training phase and the 6ST/1DT group the slowest). However, when the transfer effect was examined, all groups achieved similar levels of learning despite overall RT differences. The results from these two experiments suggest that distraction during practice had little impact on sequence learning per se.

Frensch et al. noted that this latter observation is not totally inconsistent with Curran and Keele's formulation as it can be argued that learning essentially reflects the operation of the non-attentional mechanism. However, the dual-mechanism theory predicts differential amounts of learning associated with different practice schedules in the first experiment. Because learning was measured in a single-task setting where neither mechanism was inhibited and subjects of one group received more training in the single task setting while those in the other group more training in the dual task setting, subjects who had more single-task training would be expected, according to the dual-mechanism theory, to exhibit more learning than subjects who had less such training. Obviously, the dual-mechanism theory has difficulty accommodating data where no difference in learning was observed, unless it can be shown that attentional learning can be complete within a single block of practice when the SRT task is

performed without distraction.

Frensch, et al. ran a second pair of experiments to rule out this possibility. In both, groups that received only one block of practice before learning was assessed (1DT and 1ST) were compared with groups that received 5 blocks of practice (5DT and 5ST). The two experiments differed only in that the primary stimulus followed an ambiguous sequence in the first but a unique sequence in the second. Similar performance was observed in both experiments despite sequence type difference. Of special interest was the finding that learning was not complete in a single block of practice. Specifically, subjects who received 5 blocks of practice demonstrated greater learning than those who received only one block of practice whether learning was measured with or without the secondary task. Frensch, et al. concluded that implicit sequence learning can be mediated by a single learning mechanism which is not affected by the availability of attention, although such learning may not be expressed under conditions of severe distraction which interferes with retrieval.

1.2.3 The secondary task disrupts grouping in sequence learning

The view that a secondary task reduces the amount of attention critical to the operation of the hierarchical encoding mechanism was also challenged by Stadler (1995), who argued that the secondary task impairs sequence learning because it disrupts organization of successive events during encoding. Stadler (1993) had previously demonstrated the importance of grouping in sequence learning by randomly placing a longer (2000 ms) than usual (400 ms) response-stimulus interval (RSI) either at the end of each repeating sequence or between two trials of each sequence. The

disruptive effect of the randomly occurring long RSI reduced learning significantly.¹

Although this study used the single task procedure, it suggests that the introduction of the dual task might operate in much the same way—that is, by interfering with the subjects' attempts to organize the ongoing sequence in a consistent manner.

To test this conjecture, Stadler (1995) had one group of subjects perform the SRT task with the secondary tone-counting task; a second group was required to retain a memory load of a set of letters while performing the SRT task; and a third group performed the SRT task as a single task but with inconsistent grouping cues. In addition, two other groups, who served as controls, received either the letter sets or the tones, but were told to ignore them. Stadler found that subjects who worked with the secondary tone-counting task closely resembled those with the inconsistent cues and both showed less learning than those with the separate memory load. Stadler's argument is that while memory load engages one's attention, it does not disrupt sequence grouping. In contrast, the randomly occurring long pauses place little extra demand on attentional expenditure, but they do impede consistent grouping. The functional similarity in subjects' performance in the tone-counting condition and the long pause condition invites the inference that similar processes were disrupted during learning. Incidentally, it also suggests that the secondary tone-counting task may not be an effective means to manipulate the availability of attention because its detrimental effect on sequence learning may be attributable to its disruption in

¹ In a recent study, Willingham, Greenberg and Thomas (1997) showed that inconsistent RSIs, by itself, did not disrupt sequence learning. Subjects performing the SRT task with randomly determined RSIs displayed similar learning as those with a constant RSI.

representational organization rather than a reduced attentional capacity.

The role of grouping in the SRT task had been examined earlier by Keele and Jennings (1992) in a study designed to test a connectionist model of the task based on the recurrent feedback algorithm developed by Jordan (1990). Both human subjects and a Jordan network produced similar data when trained with unique, hybrid and ambiguous sequences. Of special interest was the finding that when an ambiguous sequence was suffixed with an additional unique element (e.g., ACBCABD), both the human subjects, who performed the task under distraction, and the neural network, displayed sequence learning as if the sequence were of the unique type. Keele and Jennings argued that the salient suffix served as a grouping cue which facilitated the encoding of the sequence. The idea is further supported when a parsing mechanism was introduced in the computer model, which was implemented by resetting the state units to zero at the end of each repetition of the sequence. Consistent grouping induced by parsing soon led the neural network to produce near-perfect prediction.

As an aside here, it is worth noting that studies using the artificial grammar (AG) learning task provide additional support for the importance of organization in implicit learning (Servan-Schreiber & Anderson, 1990). During testing they presented subjects with strings of letters that either were consistent with the rules of the AG or violated them in a subtle way. When the presentations followed the natural chunks of the grammar (i.e., groups of letters with high degree of covariation were presented together) subjects performed well, reliably distinguishing well-formed from ill-formed strings. When the chunks were assembled arbitrarily, performance was dramatically reduced. Interestingly, subjects who worked with unchunked displays performed like

those who worked with chunked strings based on runs of the grammar suggesting that subjects may have used these kinds of organizational tendencies on their own.

Grouping, chunking and hierarchical encoding are related ideas. They are hypothetical processes by which complex information is believed to be encoded and corresponding mental representations formed. Stadler may be right that the secondary task and the irregular pauses both disrupt grouping in the representation of stimulus events. However, demonstrating a common consequence of different disruptive procedures does not necessarily imply that the same mechanisms are actually involved in producing the disruption. The random pauses disrupt grouping because they encourage subjects to encode the event sequence in many inconsistent ways imposed by the pauses as grouping cues. The distraction task, however, may disrupt encoding in a somewhat indirect way. To speculate, encoding a sequence with non-unique transitions involves extracting subsegmental covariations, identifying markers for these segments and then integrating local rules pertaining to successive segments of the event sequence to form higher-order conditionals. In computational terms, carrying out these processes requires considerable amounts of time and memory space in both encoding and retrieval. When the processing capacity is reduced by the secondary task, it is small wonder sequence learning, or the expression thereof, becomes more difficult.

It should be noted, however, that the disruption in encoding due to inconsistent grouping cues cannot necessarily be blamed for all the difficulties of sequence learning in the dual-task paradigm. To support this cautionary note, let me point it out that although Stadler's (1995) tone control subjects were given the same inconsistent

grouping cues, they nonetheless exhibited markedly better learning than subjects who had to deal with the tones in a way that seriously taxes attentional resources. It is likely that subjects were somehow screening out the irrelevant auditory stimulus while gradually picking up the critical regularities underlying the visual stimulus to which they have to respond. It is not unreasonable to think of the implicit learning machinery as having evolved to withstand noise in the environment. The subjects in Cleeremans' studies (Cleeremans & McClelland, 1991; Jiménez & Cleeremans, 1994) showed strong sensitivity to sequence structure even when the event sequences, which were generated by an artificial grammar, were interspersed with 15% of random elements.

1.2.4 The secondary task reduces short-term memory capacity

While Stadler (1995) contends that a secondary task impairs implicit sequence learning by disrupting consistent grouping, Frensch and his colleagues (Frensch et al., 1994; Frensch & Miner, 1994) have attributed the detrimental effect of the secondary task to a reduction in the capacity of short-term memory. This memory-based theory presupposes a rapidly decaying (500 ms or so) working memory with a limited capacity, which varies among individuals. They argue that the disruptive effect of the secondary task could derive from simply lowering the probability that pieces of information about the target sequence are active in memory at any one time thereby compromising the detection of covariation among elements.

To test this notion, Frensch and Miner (1994, Exp. 2) ran subjects with response-stimulus intervals (RSIs) of either 500 or 1500 ms. Half of each group performed the SRT task alone and half performed it with the tone-counting task. The results showed

that all subjects learned in the single task conditions, suggesting that a delay by itself may not necessarily prevent sequence learning since maintenance rehearsal can be used to bridge it (Willingham et al. [1997] also produced similar findings). When the secondary task was performed together with the primary task both groups showed less learning but the 500 ms RSI subjects were significantly better than the 1500 ms. Frensch and Miner measured subjects' digit span prior to the SRT task as an index of their short-term memory capacity. They later computed correlation coefficients between subjects' short-term memory capacity and their sequence learning and found that the correlation reached significance only in the RSI = 500 condition. Frensch and Miner argue that these results fit their short-term memory model.

The argument goes like this. In the typical SRT study performing a secondary task tends to slow down RTs by some 200-300 ms resulting in an R, on average, of about 500-600 ms. When the RT on each trial is added to an RSI of 500 ms, the interval is approaching the duration limits of short-term memory. Hence by the time a representation of the ensuing stimulus arrives in short-term memory, the preceding one has suffered substantial decay and the probability of detecting the covariations between elements of the sequence begins to diminish. In addition, the heavy demand of the secondary task provides little chance for maintenance rehearsal. This analysis also provides a novel explanation of why sequence learning is usually poorer in a dual-task setting. With an RSI of 1500 ms, the likelihood of successive elements of the sequence remaining active in short-term memory drops dramatically, severely compromising subjects' ability to detect the associative links between adjacent events. Frensch's point (see Frensch & Miner, 1994) is that a model with a single associative

learning mechanism is sufficient to account for observed data in sequence learning.

This theoretical framework is especially interesting because it makes moot the issue of the role of attention in sequence learning except insofar as diverting the subjects' attention produces the hypothesized changes in the temporal factors which compromises learning. However, the observation that short-term memory capacity correlates with sequence learning only when RSI is short (500 ms) and in the presence of a distractor suggests an interaction between timing and a capacity limit of some sort. What is intriguing about this work is that the success or failure of sequence learning was not determined by the availability or shortage of time since a longer RSI impaired rather than facilitated learning. Instead, it suggests that some critical processes related to learning may be compromised when two competing tasks, or some processes thereof, are scheduled to be performed in the same psychological time frame.

1.2.5 The secondary task adds complexity to the learning task

Virtually all the work covered heretofore has focused on the performance of the key-pressing SRT task. It is referred to as the primary task (because it is of primary interest to the researchers) while the tone-counting task was used as a distractor. As such, the "primary task" sequences have been programmed to follow certain orders depending on the researcher's intent while the "secondary task," in contrast, has been largely left to vary at the mercy of random forces except for a few weak constraints (e.g., a probabilistically defined proportion for the high- or low-pitched tone to occur). Recently, however, Schmidtke and Heuer (in press) introduced a patterned secondary task.

In their experiments a patterned tone sequence was incorporated in some conditions. They also changed the nature of the secondary task in that, rather than having subjects keep a running count of the secondary stimuli, they instructed their subjects to depress a foot pedal in response to a high-pitched tone and do nothing when the low-pitched tone was presented (what they termed the "go/no go" task). The primary visual stimulus always followed a repeating sequence of length 6. In one experiment, the performance of this group was compared with that of several other conditions. In one condition, subjects responded to a sequence of high- and low-pitched tones determined randomly. In two other conditions, one group of subjects received a 6-element tone pattern while the other received a 5-element tone pattern. There was also a single-task control group which performed the SRT task without the distractor.

As would be expected, the single-task group demonstrated the best overall performance. Taking into consideration the likelihood that the overall RTs would be affected by response conflict when subjects had to respond both to the primary visual stimulus and the secondary auditory stimulus on the same trial when the high-pitched tone was presented, an analysis was performed on RTs from the "no go" trials only. The results showed that subjects exhibited the greatest amount of learning when the primary and secondary tasks both followed a 6-element sequence, and poorer learning when the secondary stimulus sequence followed a 5-element sequence. Not surprisingly, the least learning was found in the group for which the tones were random.

These observations were taken as support for Schmidtke and Heuer's notion of

task integration. That is, sequence learning was affected by the characteristics of both the primary and the secondary stimulus. When the secondary tone was random, subjects dealt with a sequence of events in which every other event was stochastically determined. When the tone sequence followed a fixed pattern which coincides with the visual stimulus sequence in length the task essentially became one of learning a 12-element visual-auditory sequence. Although this was more difficult to learn, it could still be accomplished. The 5-element tone sequence, when presented together with the 6-element visual sequence, was shifted one element ahead of the visual sequence in each cycle. This introduced variability to the visual-auditory event pairings in a way similar to the condition in which the tone varied randomly. As the two sequences could not be integrated in a coherent way, learning was compromised.

In another experiment, Schmidtke and Heuer found that subjects' response pattern in RTs to both the primary visual task and the secondary tone task were similarly affected by changes in sequence length in the secondary task, further indicating that task integration took place in the situation when both sequences coincided.

Nevertheless, the very fact that sequence learning was substantially superior when the tone followed a 5-element sequence than when it was random suggests that learning of the visual stimulus sequence and that of the auditory stimulus sequence may have proceeded independently.

While Schmidtke and Heuer's work is provocative, it is worth noting that their "go no/go" task constitutes a major departure from the standard procedure. In one way this modification is useful since, in the standard procedure, it has never been clear whether the tone counting task directly disrupts encoding of the primary

sequence or whether it merely impacts on short-term memory. Unfortunately, however, it was not introduced systematically and it changed the procedure in ways that the findings of Frensch and his colleagues on the role of short-term memory would certainly suggest are important. Until more careful studies are carried out comparing these two variations, their findings need to be viewed as tentative.

Section summary

This section reviewed studies that attempted to tease apart a number of possible roles the secondary task may play in sequence learning in the dual-task paradigm. Cohen et al., (1990) and Curran and Keele (1993) argued that the secondary task renders a particular learning mechanism inoperable. Stadler (1995) showed that the secondary task disrupts organization of the sequence during encoding. Frensch and colleagues first suggested that the secondary task makes learning difficult by occupying a capacity-limited working memory, which lowers the activation level of sequence elements in memory (Frensch, et al., 1994). Later, they (Frensch, Lin, & Buchner, in press) presented evidence that the secondary task suppresses the expression of learning rather than disrupting learning itself. Schmidtke and Heuer (in press) contended that the secondary task simply adds complexity to the original event sequence by inserting an auditory event between every two visual/spatial elements.

I must stress its important to understand that each of these points of view can be correct to some extent without (much) conflict. The SRT task is clearly a more complex task than any of us initially appreciated. There are many ways in which performance on it could be compromised and, hence, many different processes which could be disrupted by the introduction of an attention-demanding secondary task. The

secondary task impairs sequence learning, or it appears to do so, not because it distracts attention in a simple-minded way, but because it alters the context in which learning is supposed to take place or in which it is to be expressed. Some of the data to be described later reinforce what other researchers have shown. But they also reveal a delicate timing relationship between the primary and secondary stimuli with respect to the specific effect of the secondary task that will turn out to be of some importance.

CHAPTER 2

METHODOLOGICAL CONSIDERATIONS

In the preceding chapter, I reviewed the major findings and theoretical underpinnings of sequence learning using a SRT paradigm with a secondary task. The present research was augmented on the basis of the work extensively reviewed there both conceptually and methodologically. In this chapter, I will turn to a survey of the methods and techniques developed over the past decade that were instrumental in advancing our understanding of various issues underlying implicit sequence learning. I will evaluate the merits and weaknesses of the different approaches as we go along. The discussion in this chapter is organized around three topics: (1) methods used to demonstrate learning; (2) methods used as a check of subjects' awareness; and (3) methods used to manipulate the availability of attention.

2.1 Demonstrating Learning

Like many other psychological phenomena, learning is a covert process. Although a number of pioneers are attempting to capture learning while it is going on, using sophisticated and expensive brain imaging procedures such as fMRI (see Curran, 1995 for a review), the majority of the work on implicit learning relies on techniques available in the cognitive psychologist' traditional arsenal. In other words, learning is inferred a posteriori from behavioral data.

The most mundane way to measure learning is to ask subjects a set of questions

concerning the properties of the stimulus environment. If they answer the questions correctly, it can be reasonably assumed that learning was attained. However, in the implicit learning situation, subjects normally are not specifically instructed to acquire information about the critical aspects of the stimulus environment and usually are not aware of either the processes or products of learning. Implicit learning must be demonstrated indirectly. One alternative is to record response accuracy, reported as percentage of correct responses. Performance guided by relevant knowledge is expected to show higher accuracy than that not guided by such knowledge. This measure, however, as it turned out, is not very sensitive to learning in the SRT paradigm as accuracy of key-press response soon reaches the ceiling regardless of whether the task is performed alone or in the presence of a distracting task. Most researchers record the data anyway and use it as a check to see whether subjects performed the task adequately.

In contrast, relative response efficiency has been widely used. The basic setup is to have one group of subjects perform an SRT task that follows a random sequence, and another group perform the same task with a structured sequence. Faster responses associated with the structured sequence group over time is evidence that the subjects benefitted from some knowledge of the sequence. This method can be called the *method of absolute difference*. A slightly different method based on the same logic but corrects, to some extent, for individual differences due to general response characteristics is obtained by first calculating the within-subjects RT difference between the target (usually last) block and the initial block and then finding the difference between the sequence group and the random group. Let us call this method

the method of relative difference.

The difference methods, first used by Nissen and Bullemer (1987), are simple and straightforward. However, as subsequent research has shown, they lack the level of sensitivity to capture sequence learning, especially when learning is placed in a dual task setting. It is not entirely clear why this is the case, but one possibility may be that subjects actually acquired as much sequence information when they performed the SRT task with a distractor as they did when they performed it alone. But performing two tasks simultaneously places constraints on how fast a response can be generated after the stimulus was presented. This is the well-known psychological refractory effect. Another possibility may be that the secondary task indeed interfered with sequence learning, and the data reflect a subdued learning effect.

Another method compares the slopes of response latency computed over the training blocks between the experimental and the control groups. The positive side of this method is that it takes into consideration subjects' performance in all blocks. It assumes that the RT decrement in each block over the preceding block reflects a non-specific general practice effect plus a sequence learning effect, if any. A steeper slope reflects faster improvement, thus greater learning. The slope in the random control group gives a measure of general practice effect due, essentially, to perfection of a set skills including improved motor coordination and knowledge of stimulus-response mappings. The slope in the sequence group reflects a combination of practice effect and (tacit) knowledge of sequence constraints. By factoring out the difference between the slopes, it can reveal any advantage the sequence group has over the random group (Nissen & Bullemer, 1987; Stadler, 1992). One problem with this method is that it

assumes that learning develops gradually through practice. However, this linear learning function may not reflect the reality of the learning process in this context. Indeed, it is plausible that learning may occur in stages rather than in a continuous fashion. If this is the case, the slope difference method may not be an appropriate measure of sequence learning. Empirically, as the slope is built on data over many blocks, it pools up all the variabilities, making it relatively insensitive to small effects of sequence learning.

Various kinds of difference methods discussed so far aim at finding a reliable difference between two experimental conditions. However, researchers in sequence learning in the dual-task paradigm often find themselves dealing with small effects buried in data with large variability. This often leads to accepting the null hypothesis, as might have been the case with Nissen and Bullemer's (1987) original study. One solution to this problem is to adopt a within-subjects design in what is now referred to as the transfer procedure (Cohen et al. 1990; Frensch et al., 1994; Lewicki, Czyzewska, & Hoffman, 1987). The measure of learning derived from this method has been called the (negative) *transfer effect*. The general transfer procedure starts with a training phase, in which subjects perform the SRT task for a number of blocks in which the stimulus follows a well-defined sequence. After this phase the order of the events is changed and the training sequence is replaced either by a random sequence or a novel one². Typically, an RT increment is observed after the sequence change. Since everything else remains unchanged, the impairment can be reasonably

² For the subjects, they are just performing another block of the task. Usually, they are not informed of the sequence change, and, in virtually all cases, are not aware of this change.

attributed to the sequence change.

The transfer procedure has become the standard method in recent studies of sequence learning. Apart from its higher sensitivity, it has other advantages. For instance, this method allowed Curran and Keele (1993) to measure learning at different points during a session that encompassed both single and dual task settings. It also made it possible for Frensch et al. (1994) to examine the separate effects of unique and ambiguous transitions in a hybrid sequence. One variation of this method handles the experimental sequence effect gracefully. This is achieved by having the subjects first transfer from a sequence block to one or two random blocks, and then back to a sequence block again. To find an index of learning, the first step is to average the RTs in the blocks prior to and after the sequence change. Learning is then inferred from a difference between the RT associated with the two sequence blocks and that of the random block(s)³. This provides a mechanism to counterbalance the order effect as well as to insure that the increase in RTs cannot be attributed to fatigue effect. To distinguish the two transfer methods, let us call the one with the counterbalance check the *double transfer method*, and the basic version, accordingly, the *single transfer method*. Which one should be used remains an empirical question determined by the researcher's intuition and preference. My suggestion is to use the single transfer as an index of learning but include the double transfer in the procedure. The argument is that the single transfer measure is subject to contamination of fatigue to a lesser degree, but we need the double transfer procedure

³ If more than one random block is used, as preferred by some researchers, then an averaged RT is used in calculating the difference.

to be sure that fatigue alone can not account for the RT decrement.

One common property of all methods described so far is that sequence knowledge has to be inferred indirectly from a behavioral measure, namely, an RT difference. Another cluster of methods that has been claimed to tap subjects' knowledge of the sequence more directly has also been used by a number of researchers (Howard et al. 1992; Nissen & Bullemer, 1987). These include a *recognition test* and the manual response version of a recall test dubbed the *generate test*. In the recognition test, subjects follow the stimulus in the SRT task for a number of trials as they did in the training phase. Then they are interrupted and asked if the successive events satisfy the constraints of the practiced sequence. In one variation, for example, subjects first press the keys to respond to the stimulus for a short sequence (e.g. 213), then they were to indicate, on a four-point scale, whether the sequence segment was part of the training sequence (Reed & Johnson, 1994). In the generate task, subjects are to predict the location of the next event instead of responding to an event that has already occurred. Both of these methods use accuracy of performance as a measure of sequence knowledge. They are used with the belief that knowledge that cannot be articulated by the subjects is more likely to be revealed. Note, however, that recognition and recall methods have been used as measures of explicit knowledge by investigators of implicit memory (Schacter, 1987). As one assumption of implicit learning is that the learner does not have access to knowledge acquired implicitly, some researchers use the recall or recognition tests to screen out subjects who may become aware of the sequence (e.g., Curran & Keele, 1993).

2.2 Demonstrating Awareness

Admittedly, the claim that the processes in implicit learning are truly unconscious impinges on a very touchy and controversial issue in psychology and philosophy (for a thorough treatment of the topic, see Shanks and St. John [1994] and its peer commentaries). The following is a brief survey of the methods used by researchers of sequence learning to demonstrate awareness, or the lack of it, of the presence of a repeating sequence.

Since most studies on sequence learning were conducted to investigate it as a form of implicit cognitive processing, it is essential to demonstrate that it indeed occurs in the absence of conscious awareness on the part of the subjects. It should be noted that a number of procedural technicalities have been routinely implemented for the very purpose of inducing, or inviting, learning without awareness. For example, some form of incidental procedure has typically been adopted for such a purpose. In the classic artificial grammar learning paradigm, subjects are instructed to memorize letter strings without being told that the real task is one of rule abstraction. In the sequence learning tradition, the pattern extraction task is disguised as a motor response efficiency task to a visual stimulus. Both tasks attempt to keep subjects from becoming aware of the critical processes when they are activated by the appropriate stimulus environment.

The secondary tone-counting task was introduced initially as a technique to suppress awareness (Nissen & Bullemer, 1987, p. 5). As some subjects did report knowledge about the repeating pattern when the SRT task was performed alone, it behooves the investigator to demonstrate that learning and awareness can be dissociated. It is not unreasonable to believe that having the subjects keeping track of

a randomly occurring event that is independent of the primary task will decrease the likelihood that subjects will become aware of the target pattern. However, as it was observed, learning was adversely affected in the dual task situation. Nissen and Bullemer solved this apparent anomaly by attributing the impairment to the distraction produced by the tone-counting task. They then resorted to other techniques to salvage the validity of the original claim. One such technique took advantage of special subject populations such as amnesic patients (Nissen & Bullemer, 1987) who were able to exhibit sequence learning with no recollection of the experience.

Another technique made use of the effect of some pharmaceutical substances such as scopolamine on normal subjects (Nissen, Knopman, & Schacter, 1987). Injecting healthy subjects with a small dose of scopolamine, an anticholinergic drug, produces temporary amnesia. It is believed that scopolamine "acts to block excitatory transmission via cholinergic pathways originating in the basal nucleus of Meynert and septumdiagonal band to the neocortex and hippocampus" (Knopman, 1991, p.1018). Nissen et al. (1987) found that the ability to learn the sequence was unaffected in these subjects despite the fact that many later showed no memory of the experience. A less intrusive method to ensure that subjects who show learning are indeed unaware of the experience is simply to exclude data from those who indicated awareness of the presence of the sequence before analysis was performed (Frensch et al. 1994). Unfortunately, neither of these methods are free of criticism. The pharmacological data are valid only to the extent that sensitivity to sequence information can be demonstrated in the absence of conscious recollection of how and when the relevant information is acquired. It does not, however, guarantee learning in the absence of

conscious awareness of the process. Similarly, subjects who failed to recall or recognize sequence of events after the SRT task tell us little about their conscious experience during learning.

Perhaps, the most direct way to verify if subjects acquired sequence knowledge consciously is to ask them if they noticed the presence of some regularity about the sequence of stimulus events during the experiment, and, if they did, let them describe it⁴ (e.g., Curran & Keele, 1993). If their description matches a representation of the sequence, they are said to possess conscious knowledge of it. The verbal report method, in one guise or other, has often been used by researchers as a quick-and-dirty check on subjects' awareness. The criticism leveled against the use of this method targets the fact that failing to articulate knowledge about the sequence does not necessarily mean that the knowledge was inaccessible to the subjects. There may be any number of reasons that subjects may be unable, or unwilling, to describe their knowledge of the sequence to the experimenter (Lee, 1995). As Merikle (1992) pointed out, "statements or comments indicating an absence of relevant conscious experiences may simply reflect biases introduced by either the experimental setting or the subjects' preconceived ideas concerning the value of particular types of perceptual experiences for guiding decisions" (p793). In this regard, verbal report does not qualify as a reliable and objective method for demonstrating the absence of awareness. The phenomenologists, however, would maintain that consciousness is a subjective experience after all, hence must be verified against the experiencer's verbal report

⁴ This method has been implemented either in direct communication between the subject and the experimenter, or in the form of a questionnaire in which subjects answer various questions designed to reveal their awareness of the presence of the sequence.

(see Reber, 1997).

Two other measures have also been used to probe if subjects are aware of the knowledge that guided their performance on the SRT task (Cohen. et al., 1990; Nissen & Bullemer, 1987). These are the recognition and generate method discussed in the preceding section of this chapter. Perruchet and Amorim (1992) argued that they be used as more sensitive measures of awareness provided that reference be made to the prior training experience. When the generate method was used, they found that subjects' conscious knowledge correlated highly with RTs at ending positions of sequence fragments. This led them to conclude that the two measures, that is, RTs and prediction accuracy, reflect a common database and thus are functions of the same learning processes (but see Cohen & Curran, 1993, for a rebuttal). Similar findings were reported by Jiménez and Cleeremans (1994). They also found choice RT and generation performance to be highly correlated at the global level⁵. However, they followed the recommendations of Merikle and Reingold (1991) and computed a third index which they termed an *average conditional probability* (ACP), which is the probability that an event will occur given that specific other events have been presented. After that, they calculated partial correlations of RT and generation accuracy with ACP. The analysis showed that the correlation between RTs and ACPs was still significantly greater than zero after the correlation between generation accuracies and ACPs had been partialled out suggesting that at least part of

⁵ It should be noted that the successive stimulus events in Jiménez and Cleeremans experiments comply with an artificial grammar rather than a repeating pattern. It is likely that the knowledge with respect to the behavior of the stimulus may be more complex and more abstract, thus making it more difficult to ascend into consciousness. However, the merit of the method is applicable to repeating sequence studies as well.

the performance must be accounted for by knowledge other than that which was explicit.

Granted that the generation performance reflects one's conscious knowledge about some aspects of a sequence of events, there is still evidence that some knowledge must be characterized as unconscious. Unfortunately, despite the efforts of many, questions concerning the nature of sequence knowledge still remains largely controversial. Prudence and pragmatics suggest that direct (e.g., generation) measures should be used in conjunction with the indirect (RTs) measures to assess whether the information in question is relatively available to consciousness. According to Merikle and Reingold, the only theoretical assumption that this approach requires is that the sensitivity of a direct discrimination to conscious, task-relevant information be greater than or equal to the sensitivity of a comparable indirect discrimination (Merikle & Reingold, 1991, p. 566). The implication is that insofar as an indirect measure demonstrates that the subjects' sensitivity to some relevant features of the stimulus is greater than a comparable direct measure, then a reasonable inference can be made that nonconscious processing has taken place.

2.3 Manipulating Available Attention

Although most theorists now do not equate awareness with attention, many techniques used to manipulate the availability of attention were introduced, whether announced or not, as a means of suppressing awareness. Without getting entangled in yet another controversial issue, namely, the relation between attention and consciousness, this section will only survey the methods used to manipulate the variable of attention.

By far the most popular method is the tone-counting task introduced by Nissen and Bullemer (1987). The general technique of having subjects responding to a tone while they engage in another task has been used previously and in areas other than sequence learning (Eysenck, 1982, 1995; Pashler, 1989, 1993, 1994). In the context of sequence learning studies, this technique is favored presumably because it works well as a distractor while another task is being performed. Two components of the tone-counting task make it especially demanding of attention. First, detecting the tone when it is presented entails a high level of vigilance. Second, maintaining a running count engages a substantial portion of the working memory capacity. The former requires at least a fraction of the attentional resources be allocated overhead. The latter, which involves recycling and/or updating the tone-count, also commands considerable amounts of attention right after a tone is presented. In between, subjects must process the tone perceptually and decide whether an increment of the tone-count is necessary. Pashler's (1993) work showed that, although perception of several stimuli can be accomplished in parallel with little interference, selecting a response from a set of alternatives requires the attention of the central processor, which operates in a serial fashion. As the SRT stimulus is typically presented soon after the tone, response to the primary visual stimulus (R_1) is invariably delayed, resulting in the widely-documented psychological refractory effect (Fagot & Pashler, 1992; Pashler, 1994).

As I will need to refer to stimulus and response in both the primary SRT task and the secondary tone-counting task later, let me digress briefly to define a formal notation scheme to facilitate the on-going discussion. I will let S and R stand for

stimulus and response respectively and use the subscript ₁ and ₂ to indicate if they are related to the primary or the secondary task. Thus S₁ and R₁ stand for stimulus and response in the primary task and S₂ and R₂, stimulus and response in the secondary task respectively.

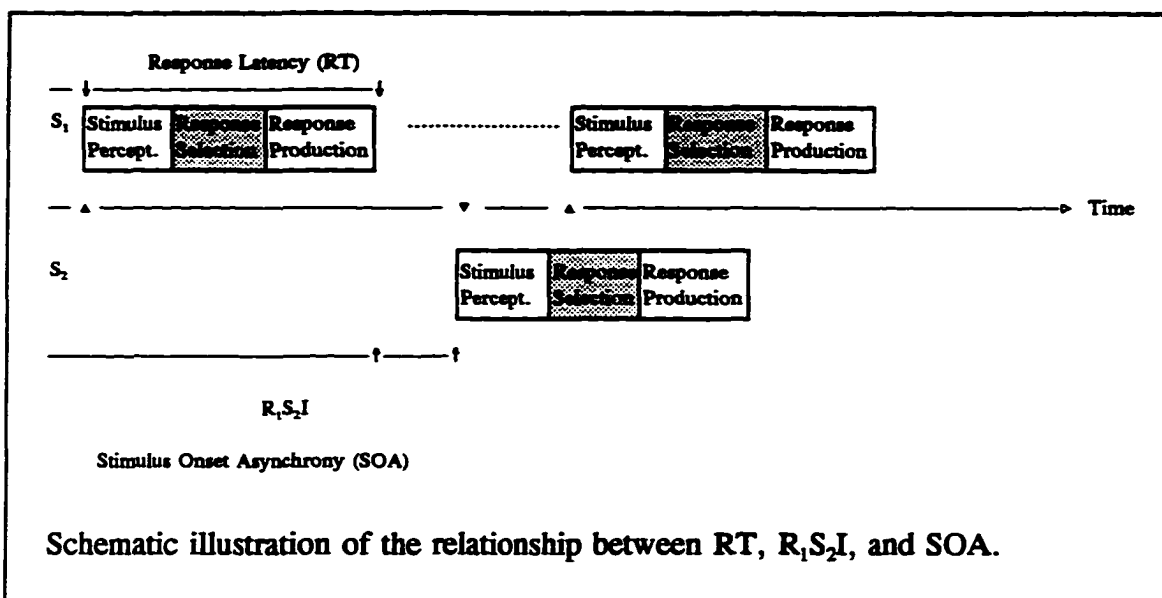


Figure 2.1 Stimulus Onset Asynchrony (SOA) in a dual-task setting: temporal relationship between the onset of the primary task and that of the secondary task.

Of interest to investigators of sequence learning is, however, not the response retardation of R₁ by the distraction task but, rather, the disruption of the secondary task on the operations to encode the sequence. Presumably, some or all processes entailed by the tone-counting task are responsible for the impairment in sequence learning. First, even if perceptual processing of the tone becomes automatized after a brief practice, deciding what to do with each tone, according to Pashler's framework, will engage the central attentional mechanism, making it momentarily unavailable to other processing. Second, keeping a running count of the target tone occupies the working memory temporarily, which, due to its limited capacity, may purge its other

contents including the representations of short runs of the primary task events. Either may prove to be detrimental to sequence learning. Note that procedurally, the tone is presented immediately, or briefly, after subjects have pressed a key in response to the primary stimulus, which, in turn, is almost immediately followed by the next primary stimulus. Both retroactive and proactive interference to sequence learning are possible. If the tone (S_2) is presented immediately after subjects' response (R_1) to the primary stimulus, the demand of the secondary task forces an attention switch. This may result in terminating the encoding process prematurely. On the other hand, if the interval between subjects' key-press (R_1) and the onset of the tone (S_2) is sufficiently long, the processes called for by the secondary task may be carried out without interfering with the primary task. However, as the response-stimulus interval (RSI) of the primary task is typically set to be so short (e.g., 200 ms), some stages of the processing on the secondary task may well overlap with the presentation, or processing, of the primary stimulus on the next trial. Due to the explicit nature of the requirement of the secondary task, it is not likely to yield more attention to the primary task other than what is necessary for detecting the presence of the stimulus in order to generate a response. Little allowance of attention is given to encode the relationship between runs of stimulus events.

In studies that involved dual task performance, oftentimes the two stimuli, S_1 and S_2 , were not presented simultaneously. The time delay between the two stimuli is customarily referred to as the *Stimulus Onset Asynchrony* (SOA) (see Fig. 1). In most sequence learning experiments (e.g., Cohen et al. 1990), the conceptual secondary task stimulus (S_2), i.e., the tone, is presented 40, 80 or 120 ms after the key-pressing

response to the primary task stimulus (S_1). As the response latency to the primary stimulus (RT) is essentially self-paced by the subjects, the delay interval between S_1 and S_2 varies from trial to trial, and from subject to subject. The secondary stimulus presentation schedule relative to subjects' key-press response has often been referred to as the response-stimulus interval (R_1S_2I). However, as I discussed in the preceding paragraph, the amount of processing overlap of the secondary and the primary tasks may play an important role in determining the quality of sequence learning. To examine this possibility, it would make better sense to shift our reference to the interval between the onset of the tone (S_2) and that of the visual stimulus (S_1) on the next trial, which I'll call the Secondary Stimulus-Primary Stimulus Onset Asynchrony (SSPSOA, or formally as S_2S_1I). Conceivably, a longer S_2S_1I (say, 160 ms) might have a lesser detrimental effect on sequence learning due to a smaller processing overlap. A shorter S_2S_1I (say, 80 ms), in contrast, may undermine sequence learning to a greater extent due to a greater amount of overlap.

There may be a situation when the two most attention-demanding stages of the two tasks are queued for processing in the central processor at the same time giving rise to a moment when demands exceed capacity. In this situation, some subjects may strive to do both, but others would opt to go easy on the counting task. Thus the former would presumably produce more accurate counts and show poorer learning whereas the latter should have less accurate counts but exhibit better learning. Surprisingly, no one has examined these effects although the R_1S_2I is typically varied in some fashion by most researchers. This potentially interesting factor has been overlooked presumably because the difference among manipulated levels was thought

to be too small to reveal any processing deficit or advantage. In addition, the variation was introduced with the specific purpose to prevent subjects from catching on to the rhythmic patterning of the visual stimulus and the tone and perceiving them as a single complex stimulus (Cohen, 1995, personal communication). In other words, the tone onset was varied to maximize its distracting effect by making it less predictable. Since the onset was determined randomly, it is impossible to isolate its effect on sequence learning.

The one study that looked into the temporal arrangement of the stimulus events and its impact on sequence learning was carried out by Frensch et al. (1994). It should be noted that the time slot they measured in this study was clocked from the onset of the primary stimulus (S_1) on the current trial to the moment a key is pressed (R_1). This is different from most of the work on sequence learning using the SRT paradigm. In the first experiment, the tone sounded either 0, 350, or 700 ms after the onset of the primary stimulus. A differential learning effect was found with these tone onset schedules. Subjects showed learning at Onset 0, those working with Onset 350 showed little or no learning, and those with Onset 700 led to two different performance patterns resulting from different response strategies. Some subjects tended to press the key prior to the tone onset; others withheld their response until the tone sounded. The former strategy yielded the best learning and the fastest overall RTs while the latter proved to be the least effective on both response efficiency and learning.

According to Frensch et al.'s estimates, pressing a key and counting the tone take approximately 450 and 300 ms, respectively when performed in isolation. They noted

that when the two stimuli are presented simultaneously, RTs are shorter than the sum of the times needed when they are performed alone. This suggests that some processes can be carried out in the same time frame without mutual interference. It is likely that the competing stimuli are processed perceptually in parallel but response selection is handled in succession. At Onset 350, the tone is presented when subjects are about to produce R_1 but may be forced to postpone it until the tone is identified. The scurry of activities lead to a surge of attentional demand. The net effect amounts to an impairment on sequence learning as well as an RT slow-down.

For subjects who receive Onset 700 and choose to respond to S_1 before they tackle the tone-counting task, they are performing the two tasks alternately. In other words, they are responding, in essence, to a single task at any time. Conversely, for subjects who withhold R_1 until they hear the tone, they may be holding the response instructions in working memory while shifting attention to the tone-counting task. As a result, subjects who use the first strategy show normal learning whereas subjects who use the second strategy exhibit very slow responses and hardly any learning at all.

Below, I will argue that sequence learning in the presence of a distractor may be accounted for by some degree of attention-sharing and/or attention-switching. The mobilization of attention is, to a certain degree, subject to voluntary control in healthy humans and perhaps other higher species in the presence of goal-directed explicit cognitive processes mediated by conscious awareness. Under optimal circumstances, several processes, including implicit ones, accomplish necessary operations by sharing the same resources. When the resources are depleted by one or more explicit

processes, the implicit processes such as sequence learning will be compromised.

It is worth noting that the secondary task in studies investigating sequence learning almost invariably involves a different (auditory) modality from the primary SRT task. The only exceptions include an unpublished study by Goschke (1992) cited in Frensch et al. (1994) and experiments carried out in our lab (Hsiao, 1995). Goschke was said to present the secondary task in the same (visual) modality as the primary task and found no sequence learning under attentional distraction. However, our data documented sequence learning with the secondary task presented either in the same or a different modality. The details will be presented later.

Other features of the secondary task are: (a) the occurrence of a target or a non-target tone is usually determined randomly; (b) the number of target tones presented in a block varies randomly except for some broad constraints (e.g., approximately 75% targets is a typical value) and (c) (as noted earlier) the tone-onset varies randomly among a set of values. Therefore, the occurrence of the target tone is essentially a stochastic event. In one sense, this property is desirable because as an effective and reliable device to control the availability of attention, it should not be affected by the strategies subjects develop and perfect in the course of the experiment. But on the other hand, these features of the secondary task may be masking psychologically interesting effects. For example, in one study, Cohen et al. (1990) varied the proportion of the target tone to manipulate the amount of residual attention available for the primary task. In one condition, the target tone occurred from 25% to 50% of the time in a block while in the other condition, it occurred from 50% to 75% of the time. Cohen et al. found that a higher frequency of the target tone was

associated with slower overall RTs. It also produced lower accuracy in the initial cycles of generation performance. Cohen. et al. argued that the higher proportion of the target tone made the task more difficult because subjects must be alert on more trials in anticipation of its occurrence, and also have to update the existing count more often. Stadler, however, suggested that the non-target tone may be actually more disruptive and consume more attention (Stadler, 1993, p.826). In our lab, we learned that some subjects developed strategies to cope with the difficulty imposed by the counting task by automatically incrementing the count by 1 on every trial regardless of its identity, only to undo it when a non-target was actually presented, or withhold counting on the next trial when the target was presented. This strategy reduces attentional expenditure, especially when target signals outnumber non-targets, as was the case in a number of studies. It also helps explain why the non-target trials appear to be more costly.

Another technique, developed by Willingham (1993), manipulates the stimulus-response mapping by varying the way the spatial layout of response keys corresponds to stimulus locations on the computer screen. In a standard "compatible" situation, subjects press the key that matches the stimulus location. In the "incompatible" condition subjects press the key one position to the right of stimulus location. In three experiments learning was found to be dramatically reduced with the incompatible mapping. Only after extensive practice were subjects in the sequence conditions able to reveal some effect of sequence learning. It appears that incompatible mapping

commands more attention until practice makes it nearly automatic⁶.

In the study of implicit sequence learning, several methodological issues have evolved. First and foremost is how to document learning. Equally important is to substantiate that what has been learned was achieved outside subjects' awareness and/or that it is beyond conscious recollection. Due to the recent debate on whether such learning requires attention, procedural techniques useful for controlling the availability of attention have also acquired considerable import. A somewhat involved discussion on these three issues will provide some guidelines for evaluating the present study as the efforts to understand them aided the conceptualization of it over the years.

⁶ More recently, Willingham (personal communication) found that the incompatible stimulus-response mapping sometimes impaired sequence learning, sometimes not. The effect appears to be somewhat unstable.

INTERLUDE: AN ONGOING SUMMARY AND SOME QUESTIONS

I have presented a general overview of the research on the role of attention in implicit sequence learning. Nissen and Bullemer (1987) arrived at the conclusion that implicit sequence learning requires attention. Thereafter, their findings have been extended time and again in subsequent investigations. In general, certain properties of the sequence, such as low predictability of component events, make it more difficult to learn. As was shown, attentional distraction from a secondary tone-counting task seemed to abolish sequence learning when the order of stimulus events could not be determined by the immediately preceding event (Cohen. et al., 1990; Nissen & Bullemer, 1987). However, later studies using different strategies showed that even events with weak constraints may be learned, although events with strong constraints are generally learned better (Frensch et al., 1994; Hsiao, 1995; Keele & Jennings, 1992). Moreover, it has become clear that learning and expression of what has been learned must be distinguished because procedures designed to assess learning may fail to detect it, or may not be able to reveal it under certain circumstances (Cohen. et al. 1990; Frensch et al., 1997). On the other hand, emphasis on the learning processes has led some researchers to suggest that procedures intended to manipulate the availability of attention disrupted groupings of events rather than reducing processing capacity (Stadler, 1995). Furthermore, some methodological refinements (such as the transfer procedure) seem to be more sensitive in detecting sequence learning under attentional distraction than others (such as the simple difference method).

The progress in the investigation of sequence learning under distraction has important implications for understanding cognitive processes and the architecture of

the mind in general. For example, the addition of a secondary task, even after extensive practice, makes responses to the primary task approximately 200 - 300 ms slower than when the task was performed alone. This suggests that there exists a bottle-neck somewhere in the processing mechanism. Conceivably, a lot of processes are being carried out in distributed and parallel fashion given the enormous amounts of information being handled by the cognitive system. However, at least some of the processes in the performance of the SRT task and the secondary tone-counting task appear to be executed sequentially. The additivity of response latency suggests that at least some stages of the two tasks are carried out in a common mechanism which has a limited processing capacity. It is likely that the attentional mechanism is capable of switching among several jobs pending processing. Performance of each task can be broken down into a number of stages such as perception, response selection and response production, as outlined by Pashler (1993). When the stimuli of two tasks are presented in such close succession that one stage of one task overlaps with that of another, the central executive switches attention between two separate tasks. Depending on the complexity (hence attentional demands) of each of the tasks and their temporal relationship, residual attention may or may not be able to sustain sequence learning.

To close in on the role of attention in sequence learning, we need to design experiments in which a number of different devices of attentional control are applied together. The availability of attention, for instance, can be influenced by having the secondary task introduced in ways such that subjects perceive it to be either more or less demanding. This will induce them to allocate more or less attention to the

distracting task accordingly. Furthermore, we can manipulate the temporal relationship between the onset of the secondary stimulus (S_2) on trial n and that of the primary stimulus (S_1) on trial $n + 1$. When different stages of the competing task are made to be performed in the same time frame, we shall be able to observe when attention is critical for sequence learning.

Admittedly, there are still many potentially interesting questions that can be answered using the SRT paradigm. While focussing on the issue of attention, the current research will address the following questions:

Question 1: Can sequence learning take place in the presence of a distractor when the sequence contains only ambiguous transitions?

Previous research has demonstrated that the structure into which the sequence of successive events was organized is critical when learning must take place in the presence of a distracting task. Specifically, if the sequence contains any unique transitions, learning can be achieved. If the events in the sequence is not predictable on the basis of the immediately preceding event, learning fails. This has lead to the conclusion that two learning mechanisms are necessary to account for sequence learning. One requires attention while the other doesn't. The present study will use ambiguous sequences only to see if the previous findings can withstand replication.

Question 2: Does sequence learning essentially involve simple association or more complicated rule abstraction?

One thing that has been under debate concerns the nature of sequence learning. Some theorists argue that sequence learning is essentially exemplar-based whereas others contend that it actually involves rule-abstraction. Since most sequence-learning studies used a repeating pattern, it is difficult to untangle the issue. It will shed light on the debate if a rule-based sequence is also used. If learning can also be demonstrated in the latter situation, it will lend substantial support to the rule-abstraction argument.

Question 3: Can sequence learning be achieved when the secondary distracting task is presented in the same modality as well as in a different modality?

The original and subsequent studies on sequence learning have always used an auditory distractor. When learning is observed in such a situation, it leaves the door open that a modality-specific attentional system is responsible for the cognitive achievement. The addition of a condition in which the secondary stimulus is presented in the same (i.e., visual) modality will expand our knowledge about the mechanisms of sequence learning under attentional distraction.

Question 4: Does sequence learning require attention?

Attention was first found necessary for sequence learning. However, recent work challenged this assertion and showed that attention may be less critical, at least under certain circumstances. In this dissertation, I will adopt several tactics to explore the effect of attention, or the lack of it, on sequence learning. The goal is to close in on

the availability of attention by combining the timing of the onset of the distracting stimulus relative to the primary stimulus, and the experimental instructions which are believed to influence the way subjects allocate attention to each of the tasks at hand.

In the chapters that follow, I will describe a series of experiments each designed to answer one or more of these questions.

CHAPTER 3

EXPERIMENT 1, EFFECT OF ATTENTION AS A SUBJECT FACTOR

This chapter addresses empirically the questions outlined above, namely, (1) whether a sequence made up of ambiguous transitions can be learned; (2) whether learning involves simple association or rule abstraction; (3) whether learning can be achieved in the same manner when the distracting secondary task is presented in the same (visual) modality as the primary task as when it is presented in a different (auditory) modality; and (4) how attention affects sequence learning. As Cohen et al. (1990) already showed that sequences made up of unique and hybrid sequences can be learned in a dual-task SRT task, only ambiguous sequences were tested in this experiment. Two types of sequences were used in the present experiment. One was the Nissen and Bullemer type of a repeating pattern; the other was generated by a rule-governed system. Departing slightly from all other studies in sequence learning, a visual secondary task was also included in addition to the usual tone-counting task. The role of attention in sequence learning will be inferred from a difference in amount of learning between subjects who paid more attention to the distractor task and those who paid less attention to it.

3.1 Experiment 1a

In a loose sense, this experiment can be viewed as a semi- replication of Cohen et al.

(1990) to the extent that the SRT task stimulus sequence followed a repeating pattern consisting of ambiguous transitions. The aim was to verify that a sequence made up only of ambiguous transitions can be learned. Apart from that, it also attempted to determine if the modality of the secondary task affected learning.

3.1.1 Methods

Subjects A total of 37 subjects (25 females, 12 males) were recruited from the Subject Pool of the Psychology Department of Brooklyn College, CUNY, where the study was conducted. These subjects' ages ranged from 17 to 45 (mean = 23.11).

The students participated in the experiment to satisfy a course requirement.

Equipment and Material A set of computer programs were developed for the study.

The main program was designed to present the stimuli and to record subjects' responses. The program set also included routines with algorithms to generate stimulus sequences and to extract summary data from raw records. The programs could be run under DOS on IBM compatible PCs. Data were collected on one of the two 486 ST 25 MHz PCs under DOS V5.0 installed in two separate cubicles.

The task is a variant of the Serial Reaction Time (SRT) task developed by Nissen and Bullemer (1987). The stimulus environment, illustrated in a diagram in Fig. 3.2 (with specifications), featured a pilot-training simulator which consisted of two concentric white circles, of which, the larger one was 30 mm in diameter, and the smaller 5 mm in diameter. A white cross originating from the common center of both circles extended from the inner circle to the outer one. The primary stimulus was represented either by one of four diamond-shaped directional arrows (\leftarrow , \rightarrow , \uparrow , \downarrow) 8 mm in length or a small solid disk (\bullet) 5 mm in diameter. The stimulus was

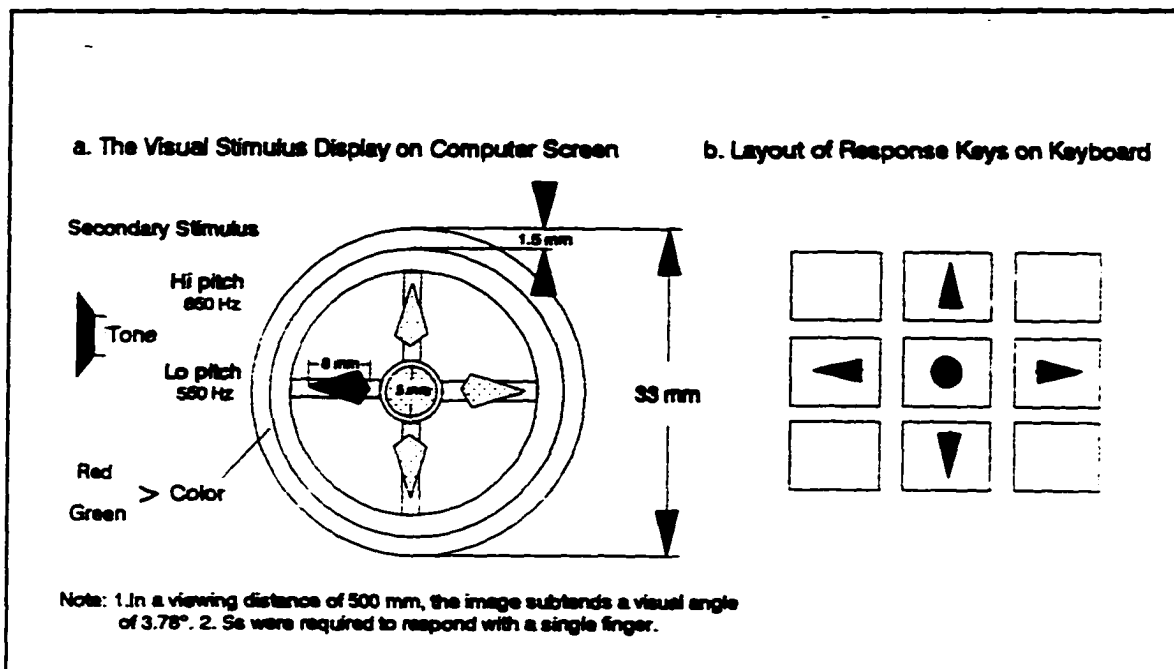


Figure 3.1 a. A diagram of the stimulus features; b. Layout of response keys on the numerical key pad.

presented in light magenta. The arrows appeared on the horizontal or vertical arms of the cross and the disk appeared in the inner circle. The arrows were used to instruct the "trainee pilot" to move the aircraft to the left, right, up or down. The disk signaled the "pilot" to keep the "aircraft" in the current position. There were two potential secondary stimuli, one visual, the other auditory, which were presented concurrently (but varied independently) regardless of whether the subjects were instructed to attend to one or the other. The visual secondary stimulus was a colored ring 1.5 mm wide immediately outside the outer circle. The auditory secondary stimulus was a tone generated by the computer's internal speaker. Each of the secondary stimuli could take on two values. The colored ring could be either red or green; the tone either high-pitched (850 Hz), or low-pitched (550 Hz). The whole visual stimulus display was centered on the computer screen 100 mm from the bottom

and subtended a visual angle of 3.78° with a viewing distance of 500 mm.

The subject was instructed to press a key in response to the primary stimulus on the numerical key pad of a standard keyboard using the keys 4, 6, 8, 2 or 5. The secondary task required the subject to keep track of the occurrence of a pre-specified target stimulus (e.g. a red ring, or a high-pitched tone). The primary stimulus code (S_1), the subject's response to each primary stimulus (R_1), and the RT associated with it were recorded on each trial by the computer as performance measures to the primary task. The codes for the two potential secondary stimuli (S_2) were also recorded regardless of which condition the subjects were assigned to. Responses to the secondary task (R_2) were numbers the subject entered as his or her counts of the number of times the target stimulus occurred in each block. In all cases, the primary stimulus (S_1) remained visible until the subject pressed the correct key in response to it. The interval between subjects' correct response (R_1) and the next S_1 was fixed at 200 ms and the secondary stimulus (S_2) was presented during this interval for a duration of 80 ms. The onset of S_2 varied randomly between 40, 80, and 120 ms relative to the subjects' key-pressing response. The temporal relations in the experiment is illustrated in Fig. 3.2.

Stimulus Sequences To determine which of the five stimulus events of the primary task is presented on any trial, two sequences were prepared, which are referred to as the Repeating Pattern (labeled Rep) and the Random Sequence (labeled Rand), respectively. The Rep sequence was used in the training phase over the first 6 blocks and in the last block (Block 8) of the testing phase. The Rand sequence, however, was only used in Block 7 in the testing phase to assess the transfer effect. So they are

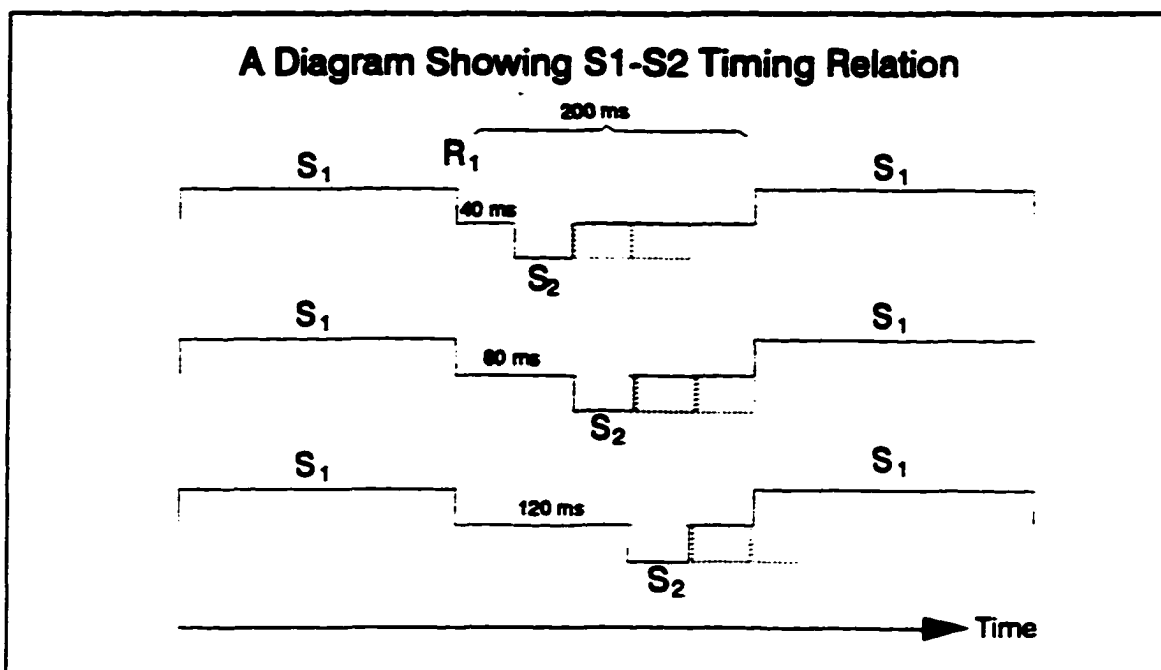


Figure 3.2 An illustration of the temporal relationship between onsets of the secondary stimulus and the primary stimulus. The boxes indicate psychological processes invoked by the secondary stimulus.

also referred to as the Training Sequence and Transfer Sequence respectively. These sequences were generated randomly by two different task-specific computer programs subject to a set of constraints: (1) by convention, no event was allowed to repeat on two consecutive trials due to the nature of the SRT task; (2) the differences in the frequency of occurrence of five distinct events (or simple-event frequencies) were kept to the minimum; and (3) the event frequencies were matched, as closely as possible, across the training and the transfer sequences (see Table 3.1 for details). The Rep sequence was obtained by first selecting the best candidate from many of the 12- element strings generated by the computer program and then repeating this string ten times to form a long sequence of 120 elements. The Rand sequence was similarly generated but the computer produced the 120 elements as a unit.

The same set of training and transfer sequences was used for all subjects. but they

Table 3.1 Sequences used in Exp. 1a and their statistical properties (simple event frequencies and 1st-order conditionals).

Training Sequence (repeating pattern):									
BDECEADCABC × 10									
Transfer Sequence (pseudo-random):									
DAEDCDCBCABE									
BDABDCECACDA									
DEAEABDECD									
BDECDBDCEBAD									
EBDBEADEDED									
BEDCDBCDABDC									
ACDADADEADCD									
CECABECBABEC									
BCADBEDCDCDC									
DBDACBACDEBC									
Simple Event Frequencies in the Sequences:									
<u>Training</u> "A" = 0.17, "B" = 0.17, "C" = 0.25, "D" = 0.25, "E" = 0.17									
<u>Transfer</u> "A" = 0.16, "B" = 0.18, "C" = 0.22, "D" = 0.28, "E" = 0.17									
Table of 1st-order Conditionals and Number of Occurrence :									
<u>Training</u>					<u>Transfer</u>				
X	BA	CA	DA 10	EA 10	X	BA 3	CA 5	DA 7	EA 4
AB 10	X	CB 9	DB	EB	AB 6	X	CB 4	DB 7	EB 4
AC	BC 10	X	DC 10	EC 10	AC 4	BC 4	X	DC 12	EC 6
AD 10	BD 10	CD 10	X	ED	AD 6	BD 8	CD 13	X	ED 6
AE	BE	CE 10	DE 10	X	AE 3	BE 6	CE 3	DE 8	X
20	20	29	30	20	19	21	25	34	20

were used only as templates to generate the actual event sequences. This was done by translating each element into one of five possible numerical codes representing a stimulus event ("A", in one version, for example, becomes "4", which represents a left-pointing arrow, ←) on the numerical key pad. Each sequence template was, in turn, translated into five different digit sequences so that stimulus locations, and hence their mapping key locations, were counterbalanced among all subjects. In addition,

both sequences qualify for the ambiguous type (Cohen et al, 1990), which means each event is followed by at least two different events. For example, the event A can be followed by either B or D (i.e., AB and AD) in the training sequence (see Table 3.1). The event B or D given A has occurred is referred to as an "A conditional" of the first order. The transition from A to B or D is ambiguous in that knowing A has occurred does not provide sufficient information to predict *with certainty* whether B or D will follow. But a quick survey of the first-order conditionals in Table 3.1 shows that some transitions in the training sequence are more ambiguous than others (e.g. D can be followed by A, C or E). Also, the training and transfer sequences, although similar in terms of simple event frequencies, turn out to be quite different from this perspective. This issue will be addressed later.

The two secondary stimuli (tone and colored ring) were generated randomly (and separately) by the computer in real-time with the constraint that the potential target stimuli (high-pitched tone and red ring) occurred 75% of the time in each block of 120 trials. Subsequently, the actual proportions of the auditory and visual target stimulus in each of the trial blocks were examined over a sample of 71 sessions (568 trial blocks). It was found that the auditory target occurred 75.20% of the time (sd = 3.87%, min. = 63.33%, max. = 85.83%) and the visual target occurred 75.26% of the time (sd = 4.00%, min. = 60.83%, max. = 85.83%).

Design and Procedure This experiment implements a 2 x 2 between-subjects design. The two independent variables were (1) secondary stimulus modality (Mod) and (2) attention available to sequence learning (Attent). Each experimental session consisted of three parts: a *training* phase, a *testing* phase and a *prediction* phase. The testing

phase, however, was a seamless continuation of, and indistinguishable from the training phase in terms of experimental procedure. Subjects were informed neither of the fact that the primary stimulus was following some pattern nor that the pattern changed some time during the session until prior to the prediction task. In the training phase, the computer program administered one of two conditions to each subject in which the secondary stimulus can be presented either as a visual or an auditory signal (Table 3.2). However, all subjects were exposed to both signals. This represents the modality (Mod) factor.

Table 3.2 Design Features in Experiment 1a.

Trial Block Setup:								
Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8	Predict
S	S	S	S	S	S	R	S	S

S = Structured Sequence; R = Pseudo-random Sequence.

Experimental Design: Primary Stimulus Sequence and Secondary Stimulus Modality:

<u>SRT Task</u>		<u>Distractor</u>
1. Rep		Auditory
2. Rep		Visual

The SRT task stimulus is presented in the visual modality in all conditions.

The second factor, which is labeled Residual Attention (Attent), was identified post hoc. Subjects were collapsed into two groups, Good Counters and Poor Counters, on the basis of their performance on the secondary task. This variable calls for some elaboration.

It was discovered during pilot studies that some subjects, for one reason or other, simply ignored the instructions to count S_2 . Such behavior defeats in many ways the

purpose of the present study. However, rather than eliminating such subjects, it seemed profitable to use their data as a "natural" control, since, presumably, these subjects were practically engaging in a single task procedure. Because of this, Attention was a subject variable not under the experimenter's control. In addition, it is a hypothetical concept arrived at through empirical inference under the assumption that the central attentional mechanism has a limited unit capacity and each of the tasks requires a portion of this capacity. Other things being equal, the amount of residual attention available to sequence encoding depends on how much attention the primary and the secondary tasks actually cost to perform. The attentional cost of the primary task, usually high at the start, presumably is reduced and remains relatively constant after some initial practice as subjects' responses become automatized. The secondary task, however, continues to be attention-demanding even after many blocks of practice. This is so partly because detection of the secondary signal remains largely stimulus-driven due to the lack of any underlying structure and partly because maintaining a running count of the target signal constantly calls for the operation of some sort of central executive. However, because the secondary task does not require an overt response until the end of each block, subjects can voluntarily control over how much attention they will allocate to the secondary task. Therefore the residual attention, or, the attentional resources actually available for encoding sequence information, will be a function of how well subjects attend to the secondary task.

Subjects were assigned to the experimental conditions randomly. This was done by drawing, from a well-shuffled stack, a coded questionnaire sheet with questions to be answered after the experiment. Each code specified one of the five versions of the

sequence set as well the modality of the secondary stimulus.

Each session began by showing the subjects a demo of the stimulus environment in which the primary stimulus occurs at different locations. During the demo the primary stimulus locations were determined randomly and its duration varied randomly between 400 and 450 ms. Then the subjects were presented the instructions which described the two concurrent tasks and spelled out what the subjects were required to do for each task. As noted earlier, the instructions put the tasks in the context of a pilot-training program and emphasized that "trainees" must respond to the "navigation direction signal" (the primary task stimulus) as fast as possible by pressing a designated key on the numeric key pad. They were asked to press the key with a single finger (of their own choice) consistently throughout the experiment⁷. At the same time, they were also required to keep track of the occurrence of a target "environmental condition alert" signal (visual or auditory, depending on the condition in which the subjects were assigned) of the secondary task. Subjects were to report the number of times the target occurred at the end of each block. It was stressed that both accuracy and speed were important in performing both tasks.

In the training and testing phases of the experiment, subjects performed the primary serial reaction time (SRT) task concurrently with the secondary target-signal-counting task for 8 blocks of 120 trials each. In the training phase (the first 6 blocks), the primary task stimulus followed a sequence of repeating pattern (Rep). In the 7th

⁷ By requiring Ss to press all keys with the same finger, the hope was to avoid complications that plagued other studies in that some fingers (e.g., the index and middle fingers) tend to produce faster RTs than others (e.g. the ring and little fingers) or some fingers are assigned to make more responses than others.

block (where the testing phase begins), the sequence was replaced by a pseudo-random sequence. In the 8th block, the repeating sequence was restored. However, as noted before, the subjects were not informed of the sequence changes. They were simply required to perform the task one block after another continuously with a short break (30 seconds) in between blocks. The increment of RTs from Block 6 to Block 7, which represents the amount of disruption in performance due to sequence change, was used as a measure of learning.

The prediction task was administered after subjects had completed 8 blocks of the SRT task. In this phase, subjects were informed that the primary stimulus actually followed a specific pattern although the actual nature of it was not disclosed. They were specifically told to discontinue the counting task. Instead, they would only be responsible for pressing a key in response to the primary stimulus on some trials, but would be prompted to predict where they thought the stimulus would appear on other trials. The stimulus followed the repeating sequence the subject received in the training phase. There were approximately 15% of prediction trials in a block of 120 trials. Whether a trial required prediction was determined randomly on the trial. On a prediction trial, the word "Predict" appeared along with an auditory alert signal. When the subject pressed a key, the corresponding stimulus appeared on the screen. However, departing from most other similar studies, no feedback was provided as to whether the prediction was correct. As was mentioned previously, this task was used as a measure of explicit knowledge about the structure of the sequence subjects may have acquired during their performance of the SRT task. Percent of correct prediction (hereafter, Pred) was used as an index of performance. The decision not to give

feedback was based on the concern that correct prediction in later part of the prediction block may be due to learning achieved earlier in the block rather than during training. This minimized the possibility that explicit learning presumed to be achieved in the training phase may be confounded with learning achieved in the course of measuring it if feedback was provided⁸.

3.1.2 Data Screening

For the primary SRT task, RTs were used to quantify performance. Typically, only RTs on trials where subjects' responses were correct were recruited in computing set-medians. In addition, subjects' key-pressing accuracy (PC_{R1}) was also obtained as a check of overall performance on the task. For the prediction task, percentage of correct prediction was computed as the measure of explicit knowledge. As anyone who is familiar with the sequence learning literature knows well, the effect of learning as assessed by an RT increment is usually rather delicate. Moreover, it is almost invariably accompanied by relatively large variability. In order to detect the effect of interest, some data cleaning is necessary. To reduced the amount of noise in the data set, two criteria, one based on overall response efficiency and the other on response accuracy, were established to determine if a subject's data would be included in the analyses. First, pilot study showed that the majority of subjects who performed the task had RTs on the primary task well under 1,000 ms. Thus any subject whose average RT over the six training blocks exceeded 1,000 ms was

⁸ Strictly speaking, we cannot rule out the possibility that (explicit) knowledge may be acquired during the prediction block by observing non-prediction trials. Knowledge about the presence of a sequence encourages rule-searching while being relieved of distraction made it all the more likely.

considered as atypical, and thus was excluded. This procedure eliminated 2 subjects. Second, data were also excluded for subjects whose primary task accuracy, averaged across all (training and testing) blocks, fell below 95%. Two more subjects were eliminated after applying this criterion.

In addition, a third screening criterion, applied on the RT measure of learning, was also adopted. This procedure would remove subjects whose learning scores (i.e., $RT_7 - RT_0$) lie beyond some reasonable range defined by two statistically determined points. The lower bound was 1.5 IQR (Inter-Quartile Range) below the 25th percentile (referred to as *the lower hinge*) and the upper bound was 1.5 IQR above the 75th percentile (similarly referred to as *the upper hinge*)⁹ (see Maoris & Well, 1991, p.14). In this experiment, the 25th and 75th percentiles were -9.5 ms and 61.5 ms, respectively. Therefore, $IQR = 61.5 - (-9.5) = 71.0$ ms. So $1.5 IQR = 106.5$ ms. An outlier would be any subject who had an implicit learning score falling beyond the range with a lower bound of -116.0 ms and an upper bound of 168.0 ms. Exploratory data analysis spotted an outlier on the implicit leaning measure in a parallel experiment (namely, Exp.1b)¹⁰. This subject showed an RT increment of 326.50 ms when the structured sequence was replaced by a pseudo-random sequence. No outlier, however, was identified in this experiment (Exp. 1a). After all data screening, 33 subjects remained.

To see if the data screening procedures were effective in reducing variability in

⁹ The procedure was recommended by Xijian Liu, a statistics professor at the Jackson State University, Mississippi, in a personal communication.

¹⁰ Experiment 1a and b were carried out at the same time with subjects assigned to each randomly. They were presented separately here for the sake of clarity.

the data, three dependent variable measures were examined, namely, primary task accuracy (PC_{R1}), the sequence learning index ($RT_7 - RT_6$) and the prediction accuracy (Pred). Before screening, average PC_{R1} was 97.96% (sd=1.63%, n = 37); after screening, it was 98.18% (sd=1.34%, n = 33). Before screening, $RT_7 - RT_6$ was 23.45 ms (sd=56.50 ms, n = 37); after screening, it is 31.83 ms (sd=52.03 ms, n = 33). Before screening, Pred = 31.80% (sd=22.79%, n = 37); after screening, Pred = 33.20% (sd=23.67%, n = 33). It seems data screening resulted in either boosting the performance measure or reducing the variance or both.

The secondary task was designed to reduce the amount of attention or processing capacity available when the primary stimulus sequence information was being encoded. Whether the device was successful depends largely on how much effort subjects actually brought into performing the target counting task. It is assumed that if the deviations of their counts (R_2) from the actual number of occurrences (S_2) were small, the distracting device was probably successful. On the other hand, if the discrepancies were large, it is likely that subjects were not paying as much attention to the task as it called for, thus leaving more capacity to handle other tasks such as encoding sequence information. A ratio of the deviation of subjects' target count from the actual number of times the target signal occurred to the actual number, which is referred to as the secondary task error rate (E_{R2}), is computed by the formula

$$E_{R2} = (R_2 - S_2 |) / S_2 \times 100\% .$$

This measure was used to assess the efficacy of the distracting task. In this formula, E, R_2 , and S_2 stand for error rate, the subject's count, and the actual number respectively. For example, if a subject reported a count of 85 times while the target

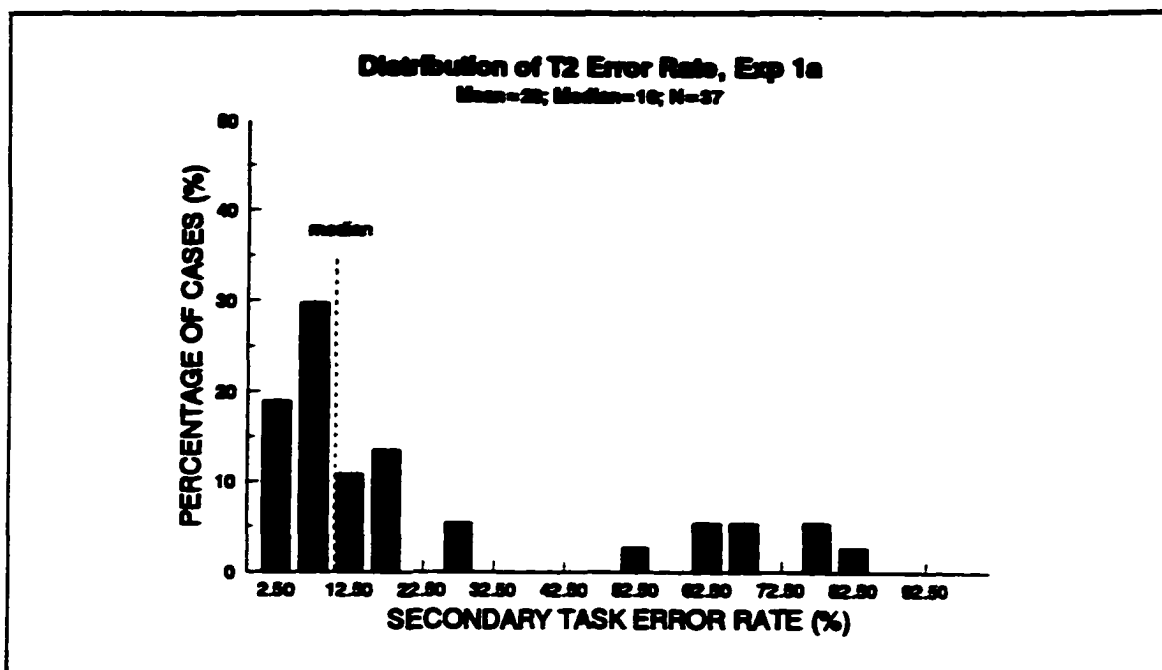


Figure 3.3 Distribution of average error rate on the secondary target-signal counting task (E_{r2}) in Experiment 1a before data were screened.

S_2 was actually presented on 90 trials in one block, the deviation will be 5 times, hence $E_{r2} = 5/90 \times 100\% = 5.5\%$. As E_{r2} may be high in the initial blocks when subjects were relatively unfamiliar with the task, an average E_{r2} was calculated for each subject without the first two blocks as an index of overall accuracy on the secondary task. Fig. 3.3 presents a distribution of E_{r2} .

The overall average E_{r2} was 22.66% (sd = 26.11%, median = 10.07%, n = 37) prior to data screening and was 20.04% (sd = 24.07%, median = 9.52%, n = 33) after screening. As is evident in the figure, approximately half of the subjects had average E_{r2} scores below 10%. The others, however, spread out in a wide range from slightly greater than 10% to over 80%. Subjects whose $E_{r2} \leq 10\%$ were considered to have performed the secondary task adequately (referred to as Good Counters hereafter), while those whose $E_{r2} > 10\%$ were considered as having performed the

task poorly (referred to as Poor Counters likewise). The assumption is that Poor Counters may have paid less attention to the secondary task, thus leaving more residual attention for encoding the sequence. Hence the average E_{22} was used to create two post hoc groups with the arbitrary cut-off point of 10% which defined the independent variable Residual Attention. The average E_{22} was 5.27% (sd=2.01%, median=5.43%, n = 17) and 35.74% (sd=26.87%, median=19.28%, n = 16) for the Good Counters and Poor Counters respectively¹¹.

3.1.3 Documenting Learning

As was noted earlier, variability has historically been a nagging problem in the SRT task. In addition to the data screening techniques described in the preceding section, various statistical devices have been used to correct for this problem, including arc-sine transformations, log transformations, trimming and /or truncating the distribution, and the use of medians. Medians have been the preferred measure in recent work. In this study, I decided to use the mean of set median RTs, where the set corresponds to one cycle of the repeating pattern, as the basic data unit because it yielded a slightly smaller variance than the simple block median¹². The mean of

¹¹ Another way to assess subjects' performance on the secondary counting task is to find the correlation between subjects counts across the trial blocks and the actual numbers. Using the median as the cut-off for categorizing subjects as Good Counters and Poor Counters produced essentially identical results in overall RTs and in learning as compared to the error rate (E_{22}) criterion. In addition, correlation of subjects counts with the true number of the intended modality is higher than that of the unintended modality, showing that most subjects were attending to the appropriate modality.

¹² Several central tendency measures were obtained on the learning effect (i.e., $RT_7 - RT_6$) in the whole sample before data screening. These include block means (Block M = 24.15 ms, sd = 76.86 ms), block medians (Block Mdn = 29.46 ms, sd = 57.40 ms), means of set medians (Set Mdn = 23.45 ms, sd = 56.50 ms), 5% trimmed means (Trim M = 26.64 ms, sd = 66.47 ms), truncated means with RTs < 1,000 ms (Trunc M =

medians for each block was obtained by finding the median RT for each set of 12 trials, excluding trials on which R_1 was in error, then computing the arithmetical mean of the medians for the ten cycles in each block of trials. However, data analysis using block means yielded qualitatively identical patterns but a generally inflated response latencies by about 100 ms. The advantage of the median-based measure is that it is resistant to the disproportionate influence of a few extreme data points (i.e., outliers) and tends to produce relatively smoother curves.

Implicit learning was indexed by an RT increment in Block 7 when the repeating-pattern was replaced by a pseudo-random sequence. In contrast, accuracy on the prediction task was used to probe subjects' explicit knowledge about the sequence. The assumption is that if subjects possess any such knowledge they should be able to predict, with better-than-chance accuracy, which event will occur given the context of preceding events. As there are five alternative choices, chance probability for our task is .20. However, it would be trivial for subjects to notice that the same event never occurred on two successive trials. Taking this into consideration, the chance level should be adjusted to .25. If the RT data actually reflect explicit knowledge, as argued by Perruchet and his colleagues (Perruchet and Amorim, 1992), rather than implicit knowledge, the two measures should be highly correlated.

3.1.4 Results and Discussion

The results will be presented in 4 sections. (1) RT measures and implicit learning;

27.40 ms, $sd = 44.12$ ms), means of log-transformed RTs (Log M = 40.61 ms, $sd = 82.03$ ms) and means of inversion-transformed RTs (Inv M = 39.43 ms, $sd = 71.73$ ms). The median based measure is opted for the sake of easy comparison with other studies in the literature and its relatively smaller variance.

(2) Pred as a measure of explicit knowledge; (3) Relation between measures of implicit and explicit learning; (4) Micro-analysis of sequence learning: simple frequencies and 1st-order conditionals.

RT measures and implicit learning

In this section, the RT data are analyzed to show two things: (1) the changes in RT with practice over the training blocks and the influence of the independent variables of attention (Attent) and the modality of the secondary task (Mod); and (2) sensitivity of subjects' RTs to a change in the sequence structure to assess if they acquired any knowledge about the sequence and the relative influence of the Attent and Mod factors.

To illustrate the effect of training on performance of the primary SRT task, RTs are plotted in Fig. 3.4 along two dimensions (namely, Attent and Mod) over the course of eight trial blocks. A 2 (Attent) \times 2 (Mod) \times 6 (Block, 1-6) ANOVA with Attent and Mod as between-subjects factors and Block as a within-subjects factor was carried out to test the effects of the factors included in the analysis. Overall, the significant main effect of Block shows that there is a general tendency for RTs to decrease with practice ($F_{(5,145)} = 25.80$, $MSe = 5901$, $p < .001$). This increase in response efficiency can be attributed to a host of factors such as familiarity with the task, improvement in motor coordination including stimulus-response correspondence, and knowledge about some properties of the sequence of stimulus events that will facilitate perceptual and/or response aspect of the task. In addition, the Good Counters tended to respond more slowly than the Poor Counters ($F_{(1,29)} = 16.78$, $MSe = 67030$, $p < .001$). This indicates that there may be a central bottle-neck

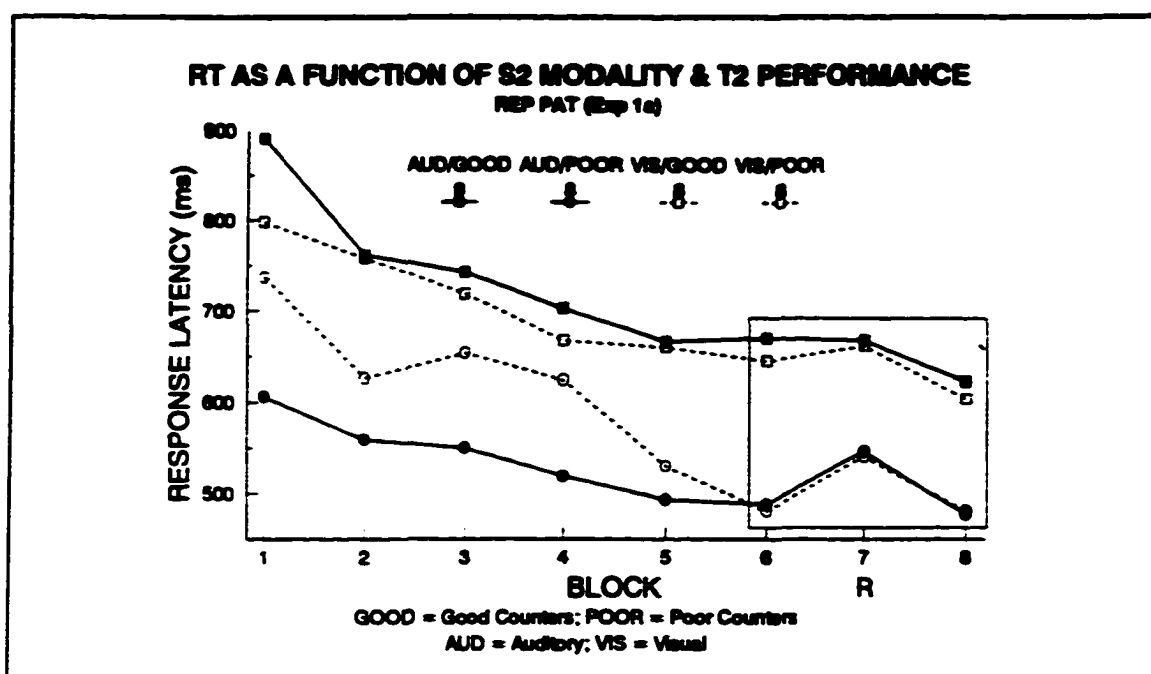


Figure 3.4 RTs over the trial blocks are plotted by how subjects attended to the secondary task and the modality of the secondary stimulus.

somewhere in the system and paying more attention to the secondary task slows down some processes of the primary task along the lines suggested by Pashler (1993).

Nevertheless, the $\text{Attent} \times \text{Block}$ interaction was not significant ($F_{(5,145)} = .87, p < .503$), suggesting that availability of attention did not affect the non-specific practice effect during training. On the other hand, the main effect of Mod was not significant ($F_{(1,29)} = .32, p = .574$), neither were the interaction of $\text{Mod} \times \text{Attent}$ ($F_{(1,29)} = 2.00, p = .169$) and that of $\text{Mod} \times \text{Block}$ ($F_{(5,145)} = .54, p = .742$), although Fig. 3.4 shows that there was some advantage of the auditory (Aud) over the Visual (Vis) modality, especially in the Poor Counters condition. There is a marginal 3-way interaction of $\text{Mod} \times \text{Attent} \times \text{Block}$ ($F_{(1,29)} = 1.95, \text{MSe} = 67030, p = .059$). This probably reflects the fact that the Poor Counters with a visual secondary task initially responded relatively slowly but eventually converged with other Poor Counters with

an auditory secondary task. It is not immediately obvious why subjects who were not paying due attention to the secondary task would need more practice to reach the same level of performance on the SRT task when the secondary task involves the visual modality than when it involves the auditory modality. Although some speculations are possible, since the effect was only marginally significant, it would be appropriate to leave it to be clarified in future studies.

The effect of learning ($RT_7 - RT_6$) with respect to *Attent* and *Mod* was tested using a 2 (*Attent*) \times 2 (*Mod*) \times 2 (*Block*, 6-7) ANOVA. The analysis produced a significant main effect on *Block*, showing an overall learning effect ($F_{(1,29)} = 16.30$, $MSe = 1092$, $p < .001$). The main effect of *Attent* was also significant ($F_{(1,29)} = 16.78$, $MSe = 22204$, $p < .001$) but the main effect of *Mod* was not ($F_{(1,29)} = .02$, $p = .776$). The significant 2-way interaction of *Attent* and *Block* ($F_{(1,29)} = 99.77$, $MSe = 1092$, $p < .004$) supports the finding that the Good Counters were more affected by the removal of the repeating pattern in Block 7. A viable interpretation is that insufficient attention to some aspects of the SRT task led to poorer learning in the Good Counters. But a visual examination of Fig. 3.4 shows that the Good Counters also appeared to be disrupted by the sequence change. The smaller decrement in RTs may be due to a ceiling effect which can not be ruled out in this study. Other interactions were not significant, which suggests that the modality of the secondary task did not affect sequence learning in any perceivable way.

Although RT was intended as the measure of the dependent variable, analyses were also performed on the primary task accuracy data (PC_{R1}) and the error rates of the secondary task (E_{R2}). First, a 2 (*Attent*) \times 2 (*Mod*) \times 6 (*Block*, 1-6) ANOVA on

PC_{R1} revealed a main effect of Mod ($F_{(1,29)} = 6.15$, $MSe = 5.74$, $p < .018$) and of Attent ($F_{(1,29)} = 17.83$, $p < .001$), but not of Block ($F_{(5,145)} = .51$, $MSe = 1.43$, $p = .771$). None of the interactions reached significance except a marginal 3-way interaction of Mod \times Attent \times Block ($F_{(5,145)} = 2.12$, $p = .057$). The two main effects show that Good Counters were generally more accurate than Poor Counters and subjects with an auditory secondary task were more accurate than those with a visual secondary task. The 3-way interaction basically reflects the fluctuation in performance over the trial blocks by the Poor Counters with an auditory secondary task. The superior performance of the Good Counters suggests that these subjects may be dispensing more efforts into performing both tasks. On the other hand, the slightly lower PC_{R1} of subjects given a visual secondary task indicates that the visual task may produce more interference on the primary task than the auditory secondary task. However, as average PC_{R1}s were rather high (greater than 95/%) among all subjects included in the final sample, and the differences between conditions, albeit consistent, were very small, these data will not be further discussed.

Subjects' error rates (E_{R2}) on the secondary task was also subjected to a 2 (Attent) \times 2 (Mod) \times 6 (Block, 1-6) ANOVA. This analysis produced a significant main effect of Attent ($F_{(1,29)} = 28.72$, $MSe = 1896$, $p < .001$) and of Block ($F_{(5,145)} = 14.32$, $MSe = 138$, $p < .001$), a 2-way interaction of Mod \times Attent ($F_{(5,145)} = 5.51$, $p = .026$) and a 3-way interaction of Mod \times Attent \times Block ($F_{(5,145)} = 3.96$, $p = .002$). All other effects, including the main effect of Mod, were not significant. The effect of Attent was expected and is not very meaningful here. The effect of Block simply reflects the fact that subjects' counts were much less accurate in the initial two

blocks. After that, they appeared to be relatively consistent. A similar analysis without the first two blocks removed this Block main effect ($F_{(3,67)} = .12$, $MSe = 66$, $p = .947$). The interaction of Mod and Attent seemed to indicate that the subjects with a visual secondary task were more accurate on their counts than those with an auditory secondary task but only when they were not paying sufficient attention to it (Poor Counters). However, this result and the lack of significance in the Mod factor may be due to a floor effect in the Good Counters whose E_{r2} s were defined to be less or equal to 10%. A 3-way ANOVA involving the same factors but on the two transfer blocks (6-7) showed only an Attent main effect. None of the other effects were significant. This indicates that performance of the secondary task was not affected by sequence change in the primary task. Moreover, the correlation between the sequence learning measure ($RT_7 - RT_6$) and the secondary task error rate (E_{r2}) was not significant ($r = .24$, $p = .182$).

Pred as a measure of explicit knowledge

Sequence knowledge was also assessed by subjects' performance on the prediction task. This task has been used widely since Nissen and Bullemer (1987) and was advocated as a check of explicit knowledge by Perruchet and others (Perruchet & Amorim, 1992). Typically, percentage of correct prediction was used as the performance measure. In this experiment, this task yielded an overall score of Pred = 26.53% ($sd = 10.96\%$, $n = 29^{13}$). Demonstrating empirically that subjects possess explicit knowledge was usually done by showing that this score is above chance.

¹³ Due to software error, Pred was recorded incorrectly for four subjects at the beginning of data collection. Data for these subjects on this variable were entered as missing in statistical analysis.

Assuming a chance level of 25%, the Pred score does not support that subjects possess any explicit sequence knowledge ($t_{(20)} = .74, p > .20$, two-tailed). However, it may be tricky to decide what is an appropriate measure of chance level. Given five distinct events, the chance probability is .20. However, in the context of the SRT task, repetition of the same event on was typically not allowed on two consecutive trials. This is because the same event presented on two successive trials would be very confusing to the viewer as the offset of the first event and the onset of the next can be easily missed. Although this becomes a property of the event sequence, and subjects may soon notice it, and react accordingly, such knowledge is not very interesting. Precluding knowledge of this piece of information, the chance probability would be .25. However, in this experiment, as in many others, the stimulus events did not occur with equal probability. Subjects noting the higher probability of some events may biased their response toward those events, achieving accuracy of greater than 25% without knowing anything about the structure of the sequence. It seems difficult to please everyone whether it is decided one way or another. However, as the interest of the present study is not to show that subjects possessed explicit knowledge, but rather the reverse, a relatively more conservative chance level may be more appropriate.

Relation between measures of implicit and explicit learning

The preceding section showed that subjects, on average, did not perform above chance level on the prediction task. It is likely that some actually did, but the effect was averaged out. If this had happened, and granted that above-chance performance reflects explicit knowledge, is this sufficient evidence to contend that the disruption

produced by sequence change must also be supported by the same explicit knowledge?

To substantiate this argument, it is necessary to demonstrate a positive correlation between the two measures of sequence knowledge, namely, Pred and $(RT_7 - RT_6)$. Subjects' scores on these two measures are plotted in Fig. 3.4. Simple Pearson correlation between the two measures was not significant ($r = -.08$, $n = 29$, $p = .683$). The partial correlation with the effect of Mod removed produced a similar result ($\beta = -.11$, $p = .683$). This suggests that the two measures likely tapped separate databases. The critical issue here is not whether performing the SRT task resulted in any explicit knowledge, but, rather, it is whether implicit learning as reflected in the RT increment due to sequence change can be accounted for by explicit knowledge presumably used in prediction. This, of course, does not rule out the possibility that predictive behavior may be actually affected, at least, partially, by

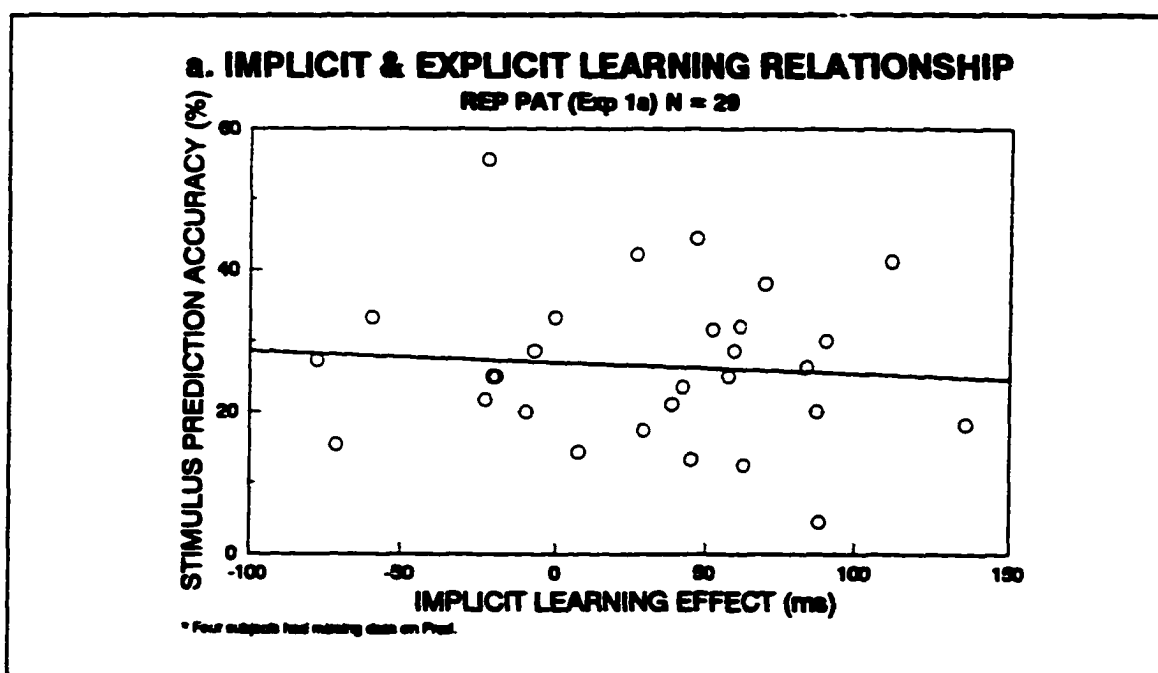


Figure 3.5 Direct (explicit, Pred) and indirect (implicit, $RT_7 - RT_6$) learning measures are plotted to demonstrate the relation between the two measure.

implicit knowledge because the task by no means entails a process-pure operation.

Micro-analysis of simple event frequencies and first-order conditionals

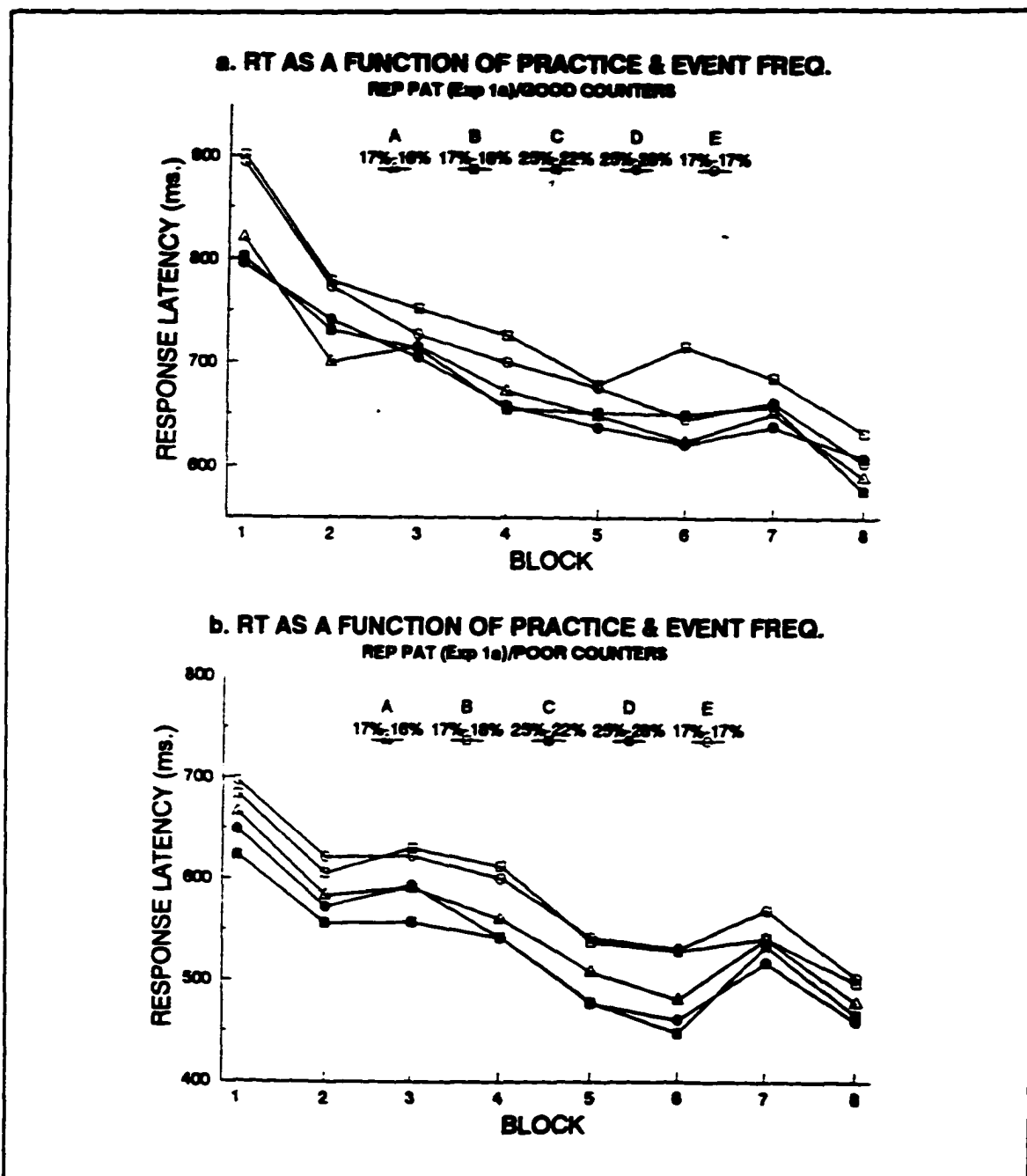


Figure 3.6 RTs associated with each of the stimulus components are plotted across the trial blocks by sequence type. Panel a for Rep sequence and Panel b for Gram sequence.

Another aspect the present experiment concerns the effect of various statistical

properties on subjects' performance of the SRT task. As was noted before, attempts were made to select a training sequence that is matched to the transfer sequence, as closely as possible, in terms of simple event frequencies (see Table 3.1). However, as noted before, some events occurred more often than others. In addition, some event pairs, called first-order conditionals, occurred in the transfer sequence but not in the training sequence. These peculiarities are worthy of some examination. They will be addressed separately in this section. Analysis of subjects' performance with respect to these properties promises to shed light on the nature of knowledge being learned. As Good Counters and Poor Counters have been found to perform the task somewhat differently, the factor *Attent* was incorporated in the analyses when appropriate.

Simple event frequencies

Two properties of the training and transfer sequences make an investigation particularly interesting. First, the five componential events were not equiprobable. Second, the training and transfer sequences were matched with respect to each of these events. From the first property, we may expect that subjects would develop differential efficiencies in their responses to these events. For example, they may become faster on trials where an event occurred more often and more slowly on trials where an event occurred less often. The mean RTs associated with each of the five events are plotted in Fig. 3.6 across the trial blocks. The data from Good Counters were displayed in Panel a and those from Poor Counters in Panel b. The numbers in percentage indicate the corresponding event frequencies in the training and transfer sequences. As is shown in the figure, events with relatively higher frequencies (e.g., both C and D occurred 25% of the time during training) tended to elicit faster

responses than events with lower frequencies (e.g., B and E occurred 17% of the time during training). A 2 (Attent) \times 5 (Event Type) \times 6 (Block, 1-6) ANOVA across the training blocks found significant main effects of Attent ($F_{(1,31)} = 18.68$, $MSe = 288674$, $p < .001$), Event Type ($F_{(4,124)} = 10.06$, $MSe = 16357$, $p < .001$) and Block ($F_{(5,155)} = 25.47$, $MSe = 28339$, $p < .001$). None of the interactions were significant ($F_s < 1$, $p > .479$). These indicate that (1) Good Counters were slower overall than Poor Counters; (2) some events (likely those with higher frequencies) were responded to faster than others; and (3) RT decreased with practice over the training blocks. However, the event A was in disconcert with this general observation. Although it belonged to the low frequency category ($f = 17\%$), it was responded to fast as higher frequency events. This suggests that event frequency may not be the only determinant in response efficiency. In addition, event frequency may be confounded with a peculiar structure of the sequence. Note that in the training sequence, B repeats itself after another event C, forming a sequel of BCB. The same happens to E, resulting in another sequel of ECE. The repeated events in such a structure have been observed to elicit longer RTs (Waugh, Fozard, Talland, & Erwin, 1973). Nevertheless, D also forms such a structure but no slowing was observed. The difference is that D occurred with a higher frequency. This issue will be addressed later.

Now let us turn to Blocks 6 and 7 to see what happened when the repeating sequence was replaced by a pseudo-random sequence. Recall that simple event frequencies were matched across the training and transfer sequences. If subjects only learned that the stimulus events occurred with differential frequencies, their

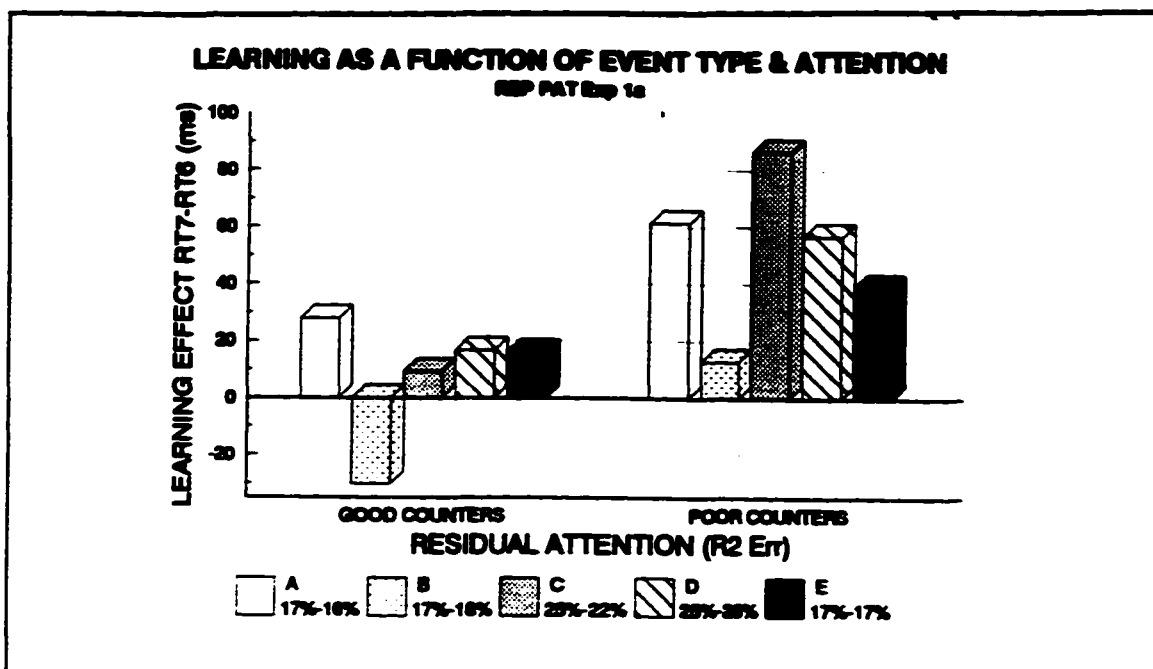


Figure 3.7 Size of learning effect associated with each of the stimulus components are shown by Good Counters (Left) and Poor Counters (Right).

performance should not be disrupted by the change of the sequence. This, however, is not true as Fig. 3.6 has shown. A 2 (Attent) \times 2 (Event Type) \times 2 (Block, 6-7) ANOVA was run on RT data in the transfer blocks. All three main effects were significant (Attent: $F_{(1,31)} = 16.06$, $MSe = 100695$, $p < .001$; Event Type: $F_{(4,124)} = 6.60$, $MSe = 5669$, $p = .001$; and Block: $F_{(1,31)} = 7.25$, $MSe = 5391$, $p < .011$). So were the Attent \times Block interaction ($F_{(1,31)} = 7.25$, $p = .011$ and that of the Event Type \times Block ($F_{(4,124)} = 3.29$, $MSe = 2603$, $p = .013$). Other interactions were not significant (for Attent \times Event Type, $F_{(4,124)} = 1.93$, $p = .110$; for the 3-way of Attent, Event Type and Block, $F_{(4,124)} = .66$, $p = .619$). From Fig. 3.7, it is clear that learning was achieved overall. It is also evident that Good Counters learned more poorly than the Poor Counters. But it is difficult to interpret the result with regard to event frequencies. Most noticeably, event B was poorly learned. However,

looking back at Fig. 3.6, the reader will notice that the poor learning may be due to an usual RT increment in Block 6 for event B. Had it followed the same course, a similar level of learning may be observed as in event E. It appears more learning was achieved in a high frequency event C, but D is more similar to the low frequency events than to C. No obvious pattern can be found concerning the effect of event frequency on sequence learning in this experiment. After all, the differences in event frequency were not very large among the component events in the sequences. Greater discrepancy in event frequencies may be necessary to reveal the effect of event frequency on learning.

First-order conditionals

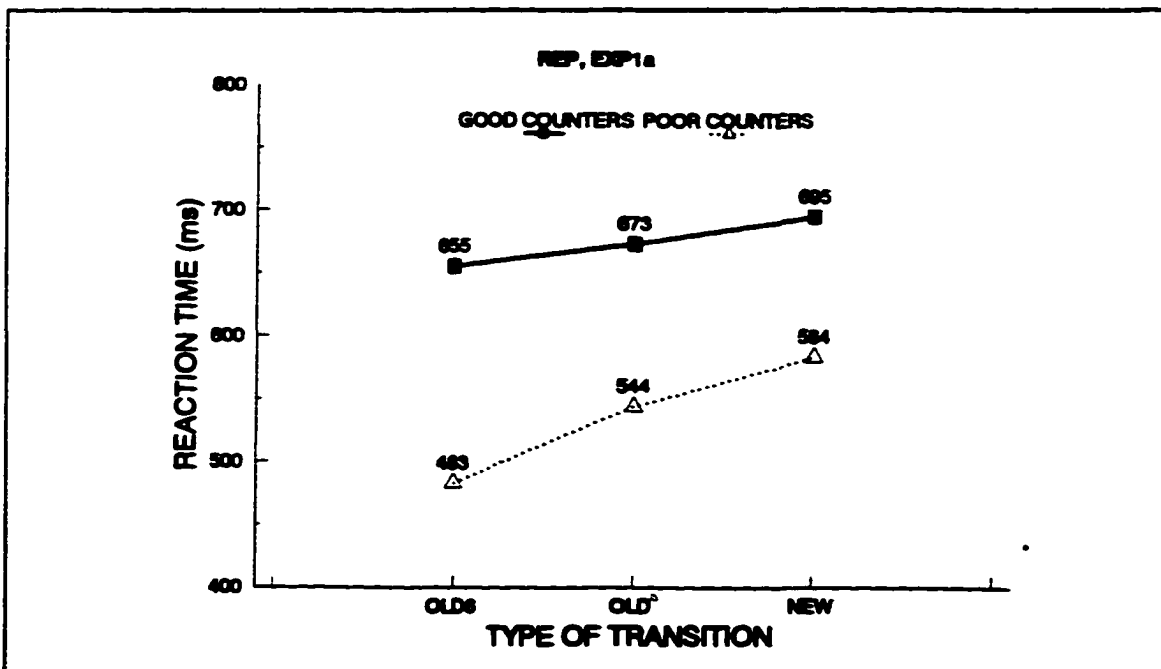


Figure 3.8 RTs plotted for first-order conditionals appeared in both the training and transfer sequence, measured in Block 6 (Old6) and Block 7 (Old) and those only appeared in the transfer sequence (New).

Two approaches were pursued to analyze the data with respect to first-order conditionals. This entails sorting subjects' RTs according to which event was

anticipated by which conditional event. For example, in the training sequence, A can be followed by B or D. So RTs on B and D trials preceded by A fall into the category of A- conditionals. Similarly, D can be followed by A, C and E. By the same token, RTs on A, C, and E trials preceded by D are categorized as the D-conditional RTs. The first approach is to compare RTs of a set of conditionals that occurred in the training sequence with corresponding ones in the transfer sequence. Since not all conditionals appeared in both sequences, five common conditional pairs (AD, BC, CD, DE and EC) were identified. RT on the trial of the second element of each pair was used as a measure of learning of that conditional dyad in that if subjects learned that when the first element was presented, the second would follow. This knowledge should facilitate their response to the anticipated second element. A 2 (Attent) \times 5 (Conditional Type) \times 2 (Block) ANOVA performed on these data revealed significant main effects of Attent ($F_{(1,31)} = 16.07$, $MSe = 112987$, $p < .001$), Type of Conditional ($F_{(4,124)} = 6.75$, $MSe = 6934$, $p < .001$) and Block ($F_{(1,31)} = 14.33$, $MSe = 14422$, $p = .001$). The interaction of Attent and Type of Conditional was not significant ($F_{(4,124)} = 1.53$, $p = .198$). But the interaction of Type of Conditional \times Block was significant ($F_{(4,124)} = 3.22$, $MSe = 7240$, $p = .015$). The interaction of Attent \times Block was also marginally significant ($F_{(1,31)} = 3.71$, $p = .061$). The 3-way interaction was not significant ($F < 1$, $p > .622$). This analysis show that subjects were disrupted by sequence change in Block 7 with respect to these trained conditionals. However, it was not uniform across all conditionals. Besides, Poor Counters probably achieved better learning of these conditionals than Good Counters.

The second approach contrasted RTs associated with two sets of conditionals, old and new, to assess the extent of learning. The old conditionals are those mentioned in the preceding paragraph, which were practiced in the transfer phase but were also present in the transfer sequence. The new conditionals are those appeared only in the transfer sequence. RTs associated with the Old conditionals could be sorted out from Block 6, which will be referred as Old6. Same could be done with the Old and New conditionals from Block 7, which will be referred to simply as Old and New. An average of each type was obtained for each subject as the basic unit of comparison. The prediction is that, if the subjects learned those five Old conditionals, their responses to them would be faster than to the New conditionals. The corresponding RTs associated with Old and New conditionals are shown in Fig. 3.8. A 2 (Attent) \times 2 (Old/New Contrast) ANOVA yielded positive evidence of learning (Old/New Contrast: $F_{(1,31)} = 16.05$, $MSe = 9996$, $p < .001$). The main effect of Attent was also significant ($F_{(1,31)} = 8.60$, $MSe = 27593$, $p < .001$). But the interaction between Attent and Old/New Contrast was not ($F_{(1,31)} = 1.51$, $p = .229$). These results suggest that subjects have learned these five dyads at least as isolated pairs and the extent of knowledge were similar in Good Counters and Poor Counters.

If the subjects learned these conditionals as no more than isolated associated pairs, there should be no disruption in response rate associated with the second member of the pair when they were embedded in a new sequence. Otherwise, the disruption must be attributed to higher-order contingencies because the larger context in which these event couplets occur was altered when a pseudo-random sequence was introduced. The hypothesis can be tested by contrasting Old6 and Old which were plotted in Fig.

3.8. A 2 (Attent) \times 2 (Conditional Type: Old6 vs Old7) ANOVA revealed significant main effects on Attent ($F_{(1,31)} = 16.73$, $MSe = 22334$, $p < .001$) and Type of Conditional ($F_{(1,31)} = 13.40$, $MSe = 1885$, $p = .001$). As is shown in Fig. 3.8, subjects were responding more slowly to the same conditionals with which they received extensive practice but were placed in a different context. The results can be interpreted to mean that learning may have extended beyond first-order conditionals. The interaction between Attent and Type of Conditionals was marginally significant ($F_{(1,31)} = 3.88$, $p = .058$), suggesting that the Poor Counters may have achieved more higher-order learning than the Good Counters. Integrating with what was found in these two analyses, it seems likely that the Good Counter's knowledge about the sequence may be more restricted than the Poor Counters. However, this needs to be demonstrated unequivocally in future studies as the interaction was only marginal.

Summary of Experiment 1a In this experiment, subjects practiced the SRT task while performing a secondary distracting task. The stimulus of the primary task followed a repeating pattern during training. It was found that subjects' manual responses to the primary stimulus was slowed down when the repeating sequence was replaced by a pseudo-random sequence. This basically replicates similar studies by Cohen et al. (1990), showing that subjects became sensitive to organizing properties of the training sequence even though they were not fore-warned of these properties nor were they explicitly instructed to look for such properties. This result argues that learning may have proceeded implicitly rather than explicitly. But this experiment also extended Cohen et al.'s (1990) findings in that a sequence with no unique transitions

can also be learned under attentional distraction. But subjects who were paying more attention to the secondary task, thus leaving less attention for encoding the sequence, showed less learning than those who may be ignoring it. In addition, the modality of the secondary task stimulus was manipulated in this experiment but it was found that it did not affect sequence learning differently. Other findings indicate that subjects' knowledge about the sequence may extend beyond simple frequencies and first-order conditionals. As the primary stimulus followed a repeating sequence in this experiment, it is likely that subjects learned the sequence as a fixed chain of events, or a few chunks thereof. It would be interesting to probe the strength of the learning mechanism by introducing a non-repeating event sequence that complies to a set of rules. Jiménez and Cleeremans (1994) demonstrated that this is possible in the absence of a secondary task when a sequence generated by an artificial grammar was used. The robust nature of the implicit learning mechanism can be further supported if it can be shown that a rule-governed non-repeating sequence can also be learned in the dual-task paradigm.

3.2 Experiment 1b

This experiment duplicate the design and procedures of Experiment 1a. The only difference is that a new sequence generated by an artificial grammar replaced the repeating pattern used in the preceding experiment.

3.2.1 Methods

Subjects A sample of 34 subjects was obtained from the same source as in Experiment 1a. Their ages ranged from 17 to 54 ($M = 23.12$, $sd = 7.78$).

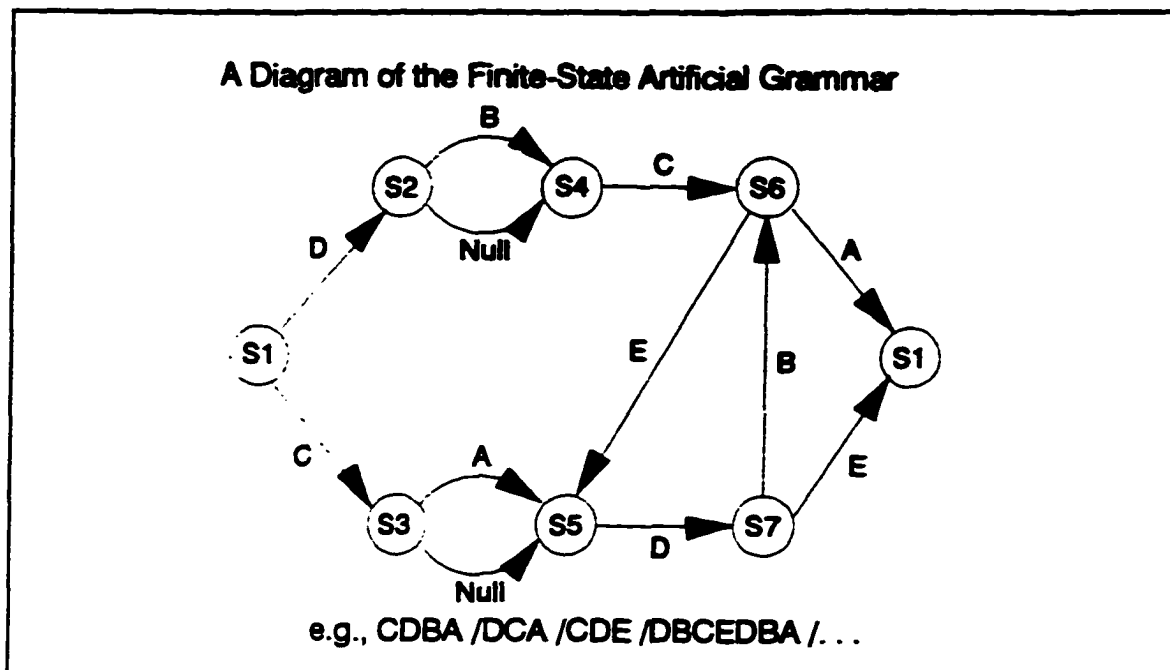


Figure 3.9 A diagram of the artificial grammar used to generate a sequence of elements to determine the order in which the SRT task stimulus is presented.

Equipment and material This experiment was run using the same equipment as described in Experiment 1a. The SRT task stimulus followed a different sequence in the training phase in this experiment. This sequence was generated by an algorithm that complies with the artificial grammar illustrated in Fig. 3.9, hence called the Gram sequence. The transfer sequence was the same one used in Exp. 1a. The statistical properties of the Gram are shown in Table 3.3.

Design and Procedures Design and procedures in Experiment 1b were identical to those used in Experiment 1a (see Table 3.3 for details).

3.2.2 Data Screening

The data were subject to the same set of screening procedures as Exp. 1a before analysis were performed. After applying the RT and PC_{R1} criteria, 30 subjects were retained. An outlier was identified on the implicit learning measure ($RT_7 - RT_6 =$

Table 3.3 Design Features in Experiment 1b.

Trial Block Setup:								
Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8	Predict
S	S	S	S	S	S	R	S	S

S = Structured Sequence; R = Pseudo-random Sequence.

Experimental Design: Primary Stimulus Sequence and Secondary Stimulus Modality:

	SRT Task	Distractor
1.	Gram	Auditory
2.	Gram	Visual

The SRT task stimulus is presented in the visual modality in all conditions.

326.50 ms for this subjects). Data from this subject were not included in further analysis. In the remaining sample, the average $PC_{R1} = 97.95$ ($sd = 1.53$). In addition, the same cut-off (10%) was used to divide the sample into Good Counters and Poor Counters. The distribution of subjects' average E_{R2} is presented in Fig. 3.10. The overall average E_{R2} in the sample was 28.10% ($sd = 30.42\%$, $media = 11.49\%$, $N = 34$). It was 4.00% ($sd = 2.58\%$, $media = 3.57$, $n = 13$) and 45.85% ($sd = 27.80\%$, $median = 45.83$, $n = 16$) for the Good Counters and Poor Counters respectively.

3.2.3 Result and Discussion

RT measures and implicit learning

The means of median RTs were used as the basic units of performance measures. Subjects' RTs are plotted across the trial blocks in Fig. 3.11 according to how well subjects performed the secondary task and the modality of the secondary task stimulus. A 2 (Attent) \times 2 (Mod) \times 6 (Block, 1-6) ANOVA was performed on RTs

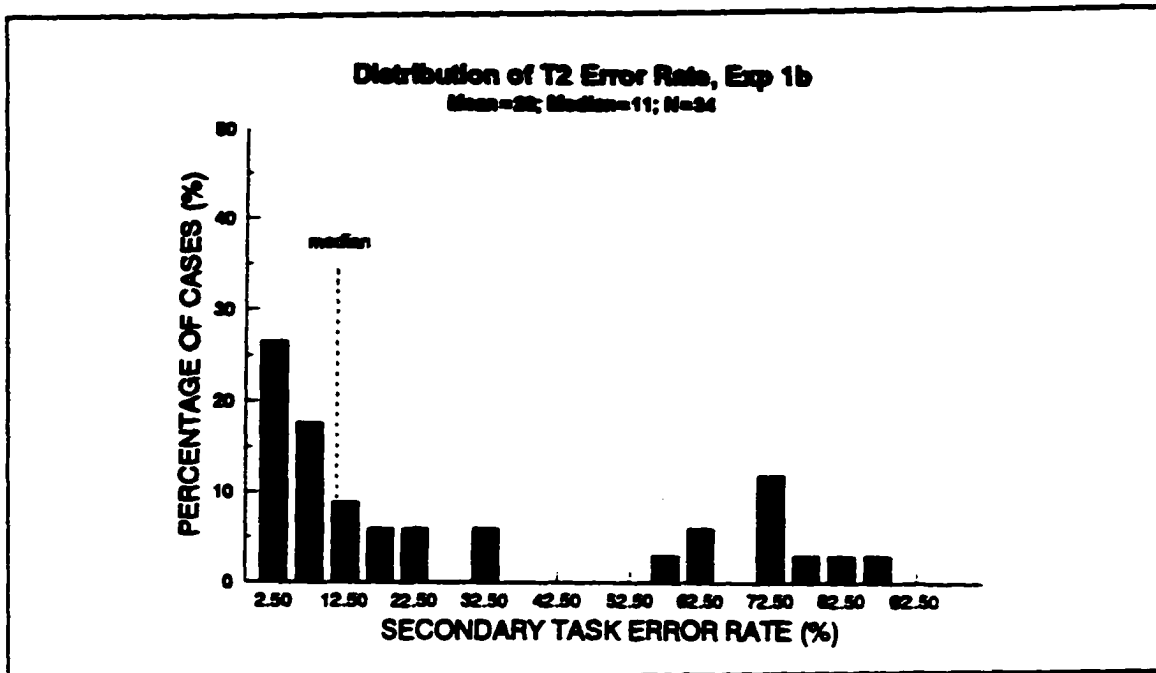


Figure 3.10 Distribution of average error rate (E_{22}) on the secondary target-signal counting task in Exp 1b before screening was applied.

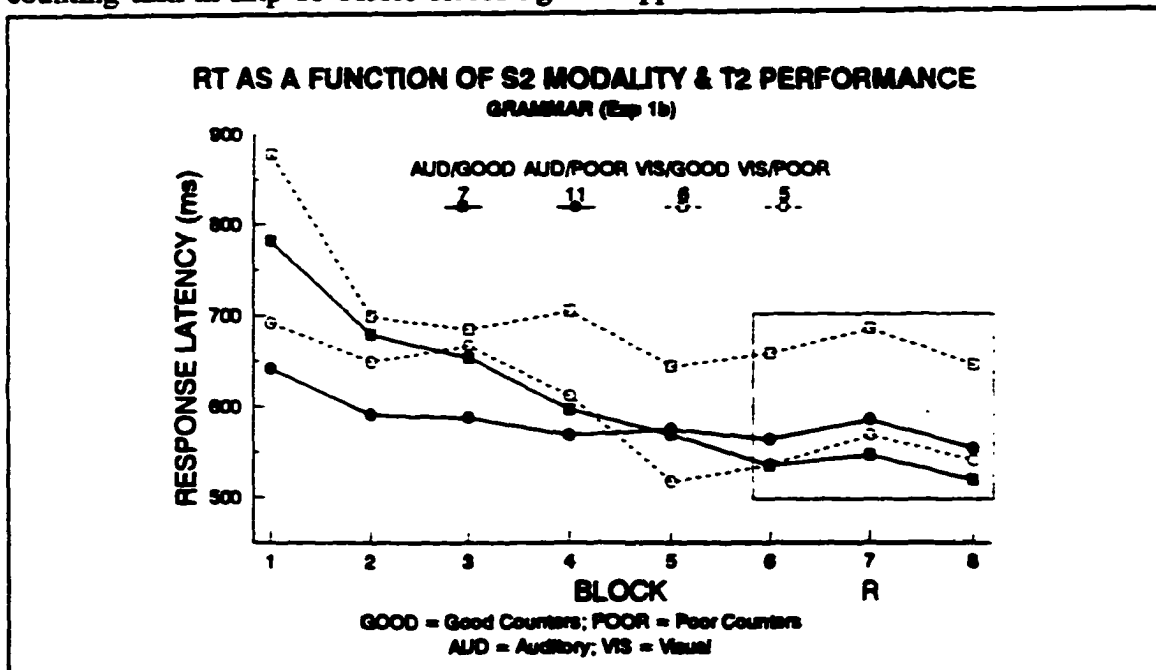


Figure 3.11 RTs over the trial blocks are plotted by how subjects attended to the secondary task and the modality of the secondary stimulus.

Attent ($F_{(1,25)} = 1.50, p = .232$). There were significant main effects on Block ($F_{(1,25)} = 6.23, MSe = 1160, p = .020$) and the Mod \times Attent interaction ($F_{(1,25)} = 4.89, p$

= .036). All other interactions were not significant ($F_s < .52, p > .483$). The Mod \times Attent interaction may be due mainly to a number of slow subjects in the group of Poor Counters who had a visual secondary stimulus. Because the sample was small ($n = 5$), the mean was greatly influenced by these subjects. Nevertheless, the main Block effect indicate that subjects were disrupted by sequence change in Block 7. The lack of the Attent \times Block interaction seems to show that Good Counters learned equally well as Poor Counters. But this may be due to a floor effect where learning was limited in both Good Counters and Poor Counters. If that is the case, the Gram sequence may be indeed more difficult to learning. However, because of the relatively small size and the differences in group size, it would be more appropriate for a future study to draw any conclusion concerning the role of attention in sequence learning when the sequence follows grammatical sequence.

Subjects' primary task accuracy (PC_{R1}) and secondary task error rates (E_{R2}) were also examined. First, a 2 (Attent) \times 2 (Mode) \times 6 (Block, 1-6) ANOVA 2 was performed on PC_{R1} and found significant main effects on both Attent ($F_{(1,25)} = 4.54, MSe = 8.95, p = .001$) and Mod ($F_{(1,25)} = 3.61, p = .027$). The main effect on Block, however, was not significant ($F_{(1,25)} = .67, MSe = 2.35, p = .645$). The interactions were not significant ($F_s < 2.36, p > .137$). For some reason, the subjects who received a visual secondary task tended to display a somewhat lower accuracy on the primary task. But this may again resulted from the small sample size of the group. 2 (Attent) \times 2 (Mode) \times 6 (Block, 6-7) ANOVA was run to see how the chance of sequence affected subjects' accuracy ont the primary task. Again, the main effect of Attent was significant ($F_{(1,25)} = 7.22, MSe = 4.69, p = .069$). The

main effect of Mod was marginally significant ($F_{(1,25)} = 3.61, p = .069$). None of the other effects were significant ($F_s = 2.75, p > .110$).

A 2 (Attent) \times 2 (Mode) \times 6 (Block, 1-6) ANOVA on the secondary task E_{22} over the training blocks showed significant main effects on Attent ($F_{(1,25)} = 31.09, MSe = 1908, p < .001$) and Block ($F_{(1,25)} = 9.70, MSe = 249, p < .001$). There was also a significant 2-way Attent \times Mod \times Block interaction. The Mod main effect was not significant ($F_{(1,25)} = 2.71, p = .112$). So were the 2-way interactions ($F_s < 1.82, p > .190$). The critical information is that the modality of the secondary task did not affect subjects performance on the task differently. Another 2 (Attent) \times 2 (Mode) \times 2 (Block, 1-2) ANOVA was performed on E_{22} . The only factor that produced a significant effect was, not surprisingly, Attent ($F_{(1,25)} = 18.90, MSe = 1038, p < .001$). Again the Mod main effect did not reach significant ($F_{(1,25)} = 2.40, p = .112$). The Block main effect was not significant either ($F_{(1,25)} = .50, MSe = 17.95, p = .486$). Neither were all the interactions ($F_s < 2.18, p > .153$). The absence of the Block main effect attests to the fact that subjects performance on the secondary task was not affected by the change of the sequence in Block 7. The important lesson of this finding is that subjects might not be aware of the sequence change. If they were, they might shift more of their attention to cope with the new sequence, causing an increase in secondary task error rate. An alternative might be to posit that secondary task performance was automatized by the time learning was measured. Given the nature of the task, this might be very unlikely.

Pred as a measure of explicit knowledge and their relationship

The implicit learning ($RT_7 - RT_6$) and explicit learning (Pred) measures are plotted

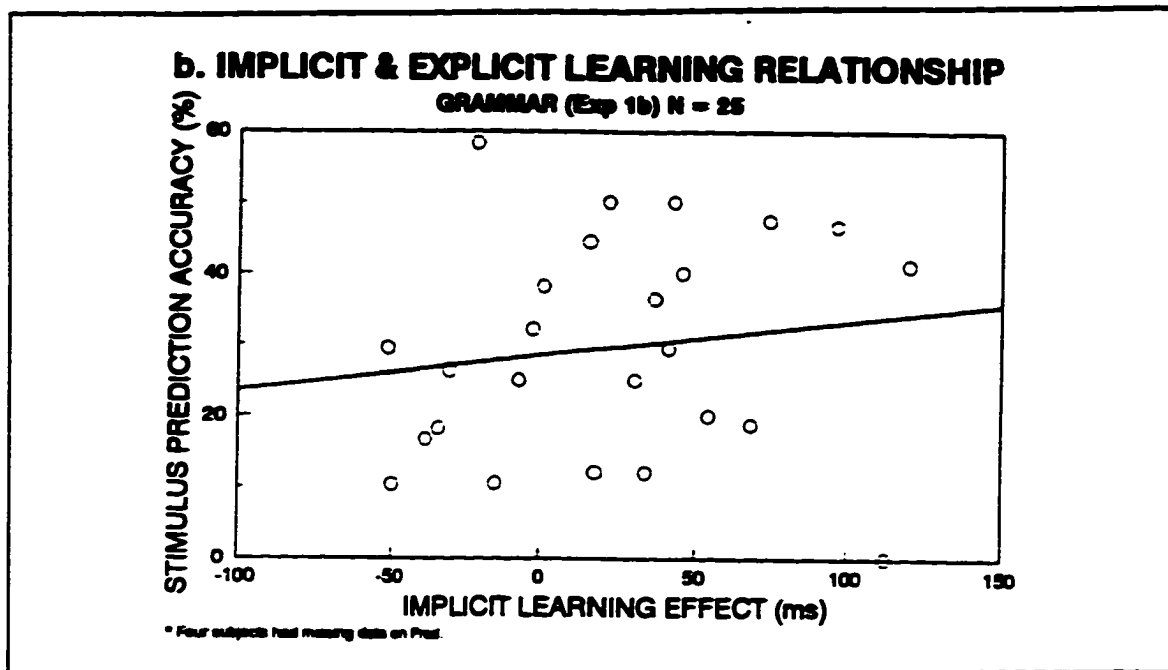


Figure 3.12 Direct (explicit, Pred) and indirect (implicit, $RT_7 - RT_6$) learning measures are plotted to demonstrate the relation between the two measure.

in Fig. 3.12). In this experiment, as in Exp. 1a, subjects' performance on the prediction task (Pred = 29.53%, $sd = 15.32\%$, $n = 25^{14}$) was not significantly different from chance ($t_{(24)} = 1.45$, $p > .20$, two-tailed), assuming a chance probability of .25. A correlation between the implicit learning measure ($RT_7 - RT_6$) and Pred yielded $r = .15$, $p = .471$ (see Fig 3.12). A partial correlation with the influence of Mod removed produced a similar result ($\beta = .22$, $p = .280$). These results lend further support to the argument that sequence knowledge was acquired implicitly and it was independent of explicitly knowledge. A discerning reader may notice that the event D occurred 29% of the time in the training sequence, which made it more salient than any other events. In addition, the DB pair occurred on a

¹⁴ As was noted in Exp. 1a, Pred was recorded incorrectly for four subjects at the beginning of data collection due to software error. Data for these subjects on this variable were entered as missing in statistical analysis.

disproportionate number of trials. Just noting these two pieces of information and guess systematically, subjects may, theoretically speaking, predict correctly well over 40% of the time. However, adjusting the conservative measure of chance level upward will not change the observation made concerning the relation between the two measures of learning.

Micro-analysis of sequence learning: simple frequencies and 1st-order conditionals

Simple event frequency

Let us first look at sequence learning associated with each of the distinct stimulus events. A 2 (Attent) \times 5 (Event Type) \times 6 (Block) ANOVA revealed a marginally significant main effect on Attent ($F_{(1,27)} = 3.64$, $MSe = 269611$, $p = .067$), a significant main effect of Event Type ($F_{(4,108)} = 4.28$, $MSe = 18357$, $p = .003$) and one of Block ($F_{(5,135)} = 11.32$, $MSe = 36413$, $p < .001$). As is shown in Fig. 3.14, subjects' RTs decreased with practice. Good Counters tended to respond not as fast as the Poor Counters. Some events (e.g., B and D) may have yielded faster RTs than others (e.g., A, C and E). An analysis on the transfer effect produced significant main effects of Event Type ($F_{(4,108)} = 7.30$, $MSe = 4666$, $p < .001$) and Block ($F_{(1,27)} = 20.17$, $MSe = 4148$, $p < .001$), but not of Attent ($F_{(1,27)} = .24$, $MSe = 79925$, $p = .630$). This suggests that the differences in the amount of learning achieved with different events can not be accounted for by the availability of attention. Ignoring the secondary task (at least some of the time) allows subjects to respond to the primary task faster. But there is little indication suggesting learning can be traced simply to those events with a higher frequency of occurrence.

First-order conditionals

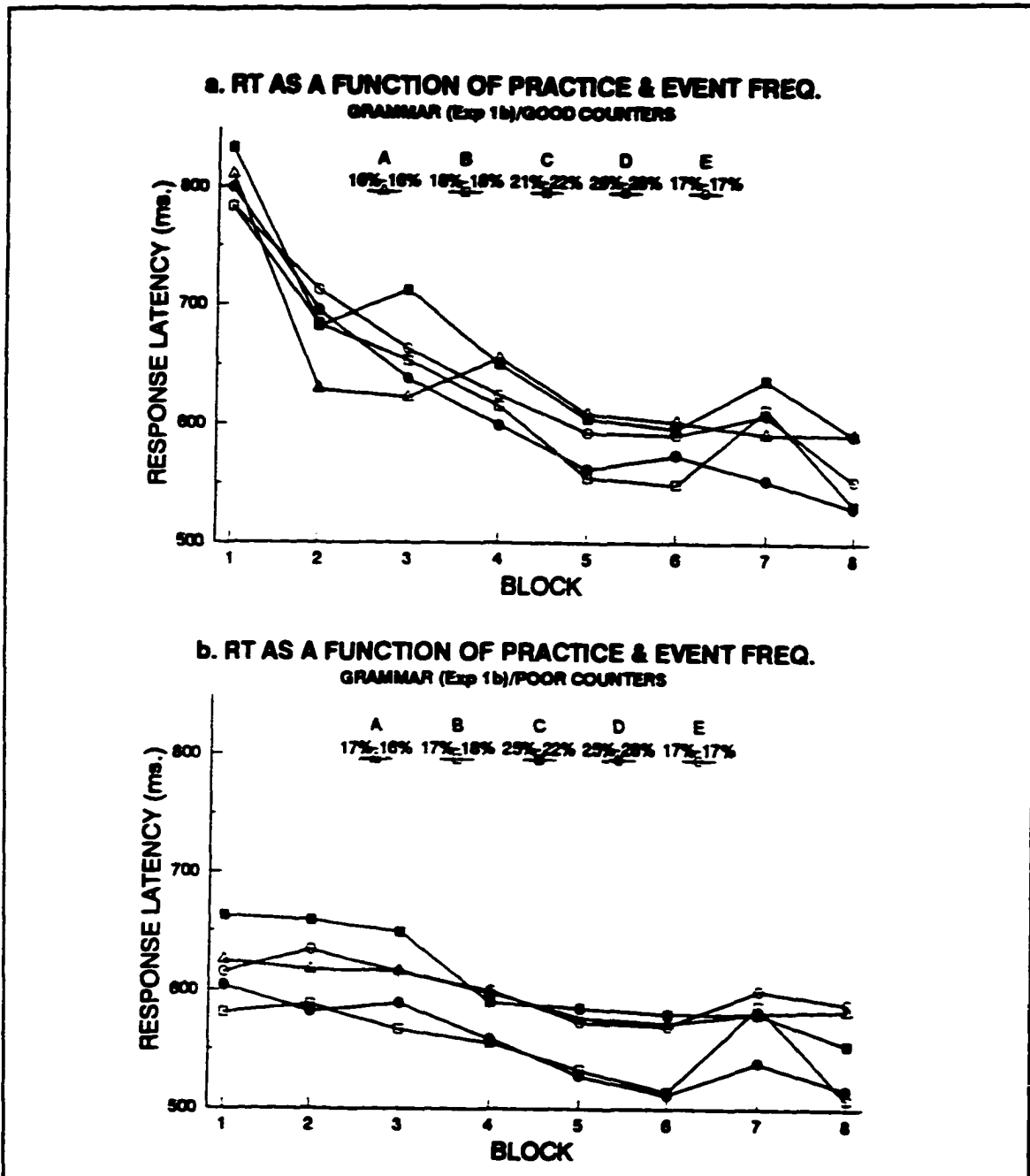


Figure 3.13 RTs associated with each of the stimulus components are plotted across trial blocks.

One event is peculiar, however. It is observed that event B was better learned than all other events by both Good Counters and Poor Counters (see Fig. 3.13). Why this particular event was easy to learn? When the training sequence for Gram was

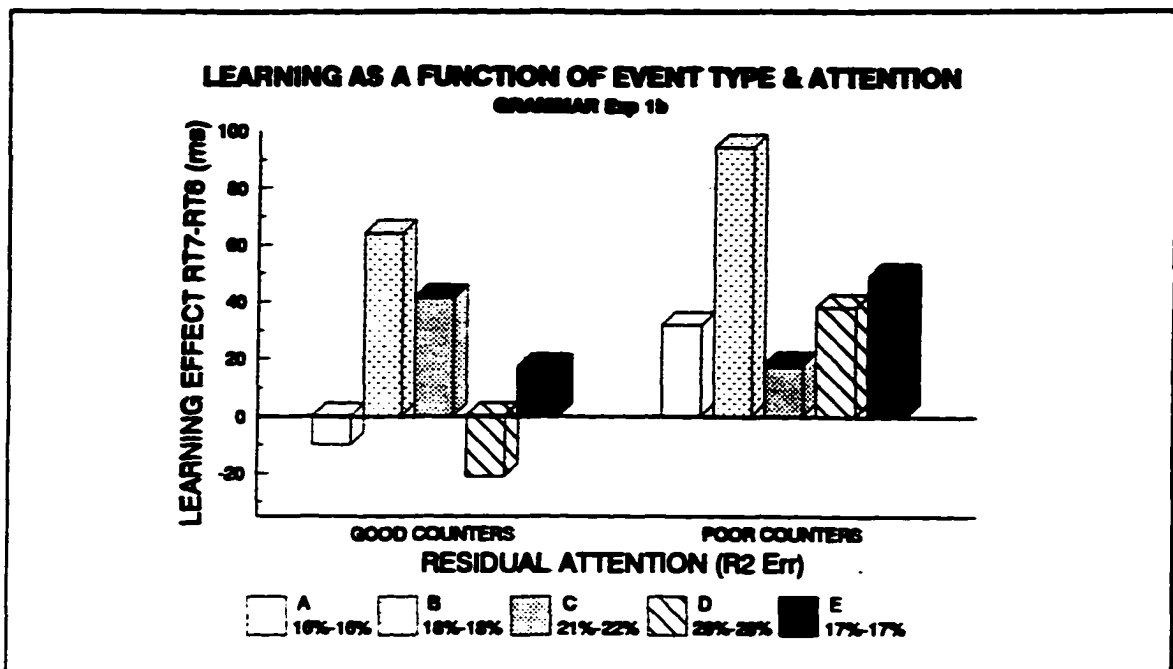


Figure 3.14 Learning effects associated with each of the stimulus components are shown by Good Counters and Poor Counters.

examined , it was found that the event combination DB occurred 21 times in a block of 120 trials (see Table 3.1). In addition, although D can be followed by two other event, C and E, B can only be preceded by D. This suggests that event frequency may be an important factor in learning at a higher level. It then becomes interesting to explore event contingencies in the sequence. It is likely that all events that precede a particular event have some impact on determining its ultimate occurrence in the chain of events. A natural starting point is to look at the event that immediately precedes it. The contingency of an event upon its preceding event, referred to as first-order conditional (Reed & Johnson, 1994), could be picked up by subjects during training. It becomes facilitative in subjects' manual responses. However, when this contingency is significantly weakened, or even disappears entirely, during transfer, a disruption should be expected. This is precisely the case in this experiment where DB transition

was absent in the transfer sequence.

As in Experiment 1a, the first-order conditionals were examined in two ways. The first approach sought to investigate if these conditionals were learned, and if the analysis showed they were, whether some of them were learned better than others. In this approach, a set of conditionals (AD, BC, CD, DE and EC) that appeared in both the training and transfer sequences were identified. Parenthetically, this is the same set entered in analysis in Exp. 1a. Appropriate RTs were sorted out as the dependent variable. A 2 (Attent) \times 5 (Type of Conditional) \times 2 (Block, 6-7) ANOVA of these RTs yielded a significant main effect of Type of Conditional ($F_{(4,100)} = 5.21$, $MSe = 19879$, $p = .001$) but not of Attent ($F_{(1,27)} = 1.49$, $MSe = 123614$, $p = .233$) nor of Block ($F_{(1,27)} = .09$, $MSe = 18057$, $p = .770$). The interactions were not significant either ($F_s < 1.55$, $p > .195$). It seems that we have little evidence that these conditionals were learned in this experiment. As the training sequence used in this experiment was generated by an artificial grammar rather than a repeating pattern, the failure to show learning of these first-order conditionals suggests that they may be more difficult to learn when embedded in a more variable context..

The second approach looked at conditionals as a group. The set of conditionals that appeared in both the training sequence as well as in the transfer sequence were identified and categorized as old conditionals. Those that only occurred in the transfer sequence were thus called new conditionals. RTs associated with the old conditionals extracted from Block 6 are labeled Old6 and those from Block 7 simply as Old. RTs associated with those new conditionals are labeled New. These data are plotted in Fig. 3.15. First, a 2 (Attent) \times 2 (Old/New Contrast) ANOVA was performed on the RTs

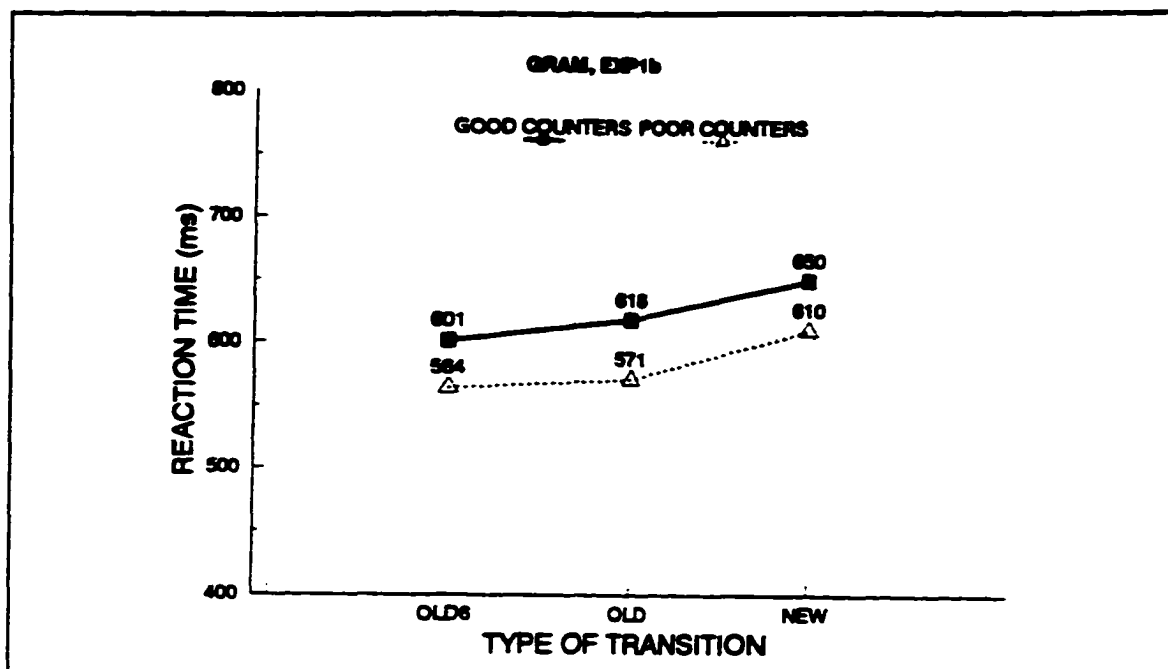


Figure 3.15 RTs plotted for first-order conditionals appeared in both the training and transfer sequence, measured in Block 6 (Old6) and Block 7 (Old) and those only appeared in the transfer sequence (New).

associated with Old and New conditionals in the transfer block. It found a significant difference in RT between Old and New conditionals within the transfer block (for the Old/New Contrast, $F_{(1,27)} = 9.20$, $MSe = 1972$, $p = .005$), indicating that the Old conditionals had an advantage over the New ones. The main effect of Attent was not significant ($F_{(1,27)} = 1.24$, $MSe = 21472$, $p = .275$). Further more, the Attent \times Old/New Contrast was not significance ($F_{(1,27)} = .09$, $p = .760$). This analysis show that subjects were able to learn first-order conditionals as associated pairs. The second 2-way ANOVA with Attent and Old6/Old as the factors found no effect of these factors (Attent: $F_{(1,27)} = 1.19$, $MSe = 21174$, $p = .284$; Old6/Old Contrast: $F_{(1,27)} = 1.44$, $MSe = 1455$, $p = .241$) or of the interaction between these factor ($F < 1$, $p = .640$). Together, these results indicate that the subjects in this experiment probably did not learn much about the sequence beyond first-order conditionals. In contrast to

what was found with the Rep sequence in Experiment 1a, this, again, is evidence that higher-order information in the Gram sequence may be more difficult to learn.

3.3 Summary and Chapter Conclusions

These experiments were designed and conducted to investigate whether attention is necessary in implicit sequence learning in a dual-task performance paradigm.

Specifically, the aim was to put Cohen et al.'s (1990) argument about attention and hierarchical encoding to a strong empirical test.

The dispute here is not on the mechanisms that are appropriate for encoding unique and non-unique transitions. One can appreciate that sequences without a unique transition are likely to be more difficult to encode. However, it appears to be an oversimplification to state that one encoding strategy requires attention while the other does not in an all-or-none fashion. Some studies (Frensch & Miner, 1994; Frensch et al., 1994) have shown that ambiguous transitions embedded in a hybrid sequence can be learned. The present experiment was designed to test the possibility that sequences made entirely of non-unique transitions can also be learned. However, while the Gram sequence used in Exp. 1b consists of non-unique transitions, it also has transitions that are probabilistically determined. Analyses at simple event frequency and first-order conditional levels revealed that the Gram sequence may be more difficult to encode than the Rep sequence (include results here). However, this may not necessarily mean that the Gram sequence can not be learned, but rather it only showed that learning was not observed under the circumstances. It is likely that given more practice, such a sequence might well be learned.

Secondary task error rate, E_{r2} , was used as a post hoc grouping variable under the

assumption that adequate performance of the secondary task requires subjects to allocate sufficient attentional resources to the task. Similarly, poor performance indicates that subjects were not engaging themselves in the secondary task. The inclusion of the secondary task as a distractor in order to manipulate attention available to sequence learning is successful only insofar as subjects commit themselves to perform the task to the best of their ability. Taking advantage of how subjects took to the secondary task as a natural way of allotting attention, this study showed that attention interacted with sequence learning when the SRT task followed a repeating pattern. When the task followed a non-repeating sequence generated by an artificial grammar, the advantage of having more attention available did not translate into better learning. A subdued level of learning suggests that a grammatical sequence may be more difficult to learn.

Another dimension that was manipulated in this experiment is the modality in which the secondary stimulus was presented. Since the primary stimulus was presented through the visual modality, a secondary stimulus in the same modality may produce greater interference than one in a different modality. The former condition may require a greater attentional expenditure for the performance of the two tasks, leaving less attentional resources available to encode structural information of the sequence. However, the experiments produced no evidence supporting this expectation. Nevertheless, subjects appeared to be somewhat slower in the training phase when the secondary stimulus was visual. But the results were not very consistent in Experiment 1a and 1b. More work is needed to clarify the precise effect of stimulus modality.

This experiment yielded a number of important findings. First, it demonstrated that subjects are able to acquire sequence knowledge in the presence of a distracting task even when the sequence consists entirely of non-unique transitions. Second, the (tacit) knowledge about the underlying structure of the sequence that the subjects acquired extends beyond simple frequency of component events. Third, it showed that the implicit learning mechanism appears to be robust in the face of variability in the structure of a sequence of events as learning was demonstrated when the stimulus followed a non-repeating sequence generated by an artificial grammar. Fourth, it seems that subjects were able to acquire sequence knowledge with similar efficiency regardless of whether the secondary stimulus is presented in the same modality as the primary stimulus or in a different modality. Fifth, subjects whose accuracy on the secondary task was relatively poor showed not only markedly faster RTs on the primary task, but also greater learning than those of subjects whose accuracy was higher. On the one hand, this indicates that performing the two tasks involves some common resources. On the other, it suggests that availability of attention is conducive to sequence learning.

Apart from these major findings, it was also observed that the modality of the secondary task stimulus contributed to the efficiency with which the primary task was performed such that RTs were faster when the secondary task involved the auditory (and a different) modality than when it involved the visual (and the same) modality. In addition, analyses on the data showed that subjects' prediction of stimulus locations was virtually at chance level. Moreover, this measure, which has been used as an explicit knowledge indicator, does not correlate significantly with the implicit learning

measure. This indicates that they are not likely to be different indices of the same database. The dissociation undermines the argument that the disruption due to sequence change reflects subjects' conscious knowledge of some aspects of the event sequence.

CHAPTER 4

EXPERIMENT 2, EFFECT OF ATTENTION AS A FUNCTION OF TASK DEMAND

Several issues were explored in the previous chapter. The results of Exps 1a and 1b showed that event sequences of the ambiguous type can be learned even when attention was diverted with the modality of the secondary task stimulus making little or no difference on sequence learning differentially. In addition, learning was observed with an artificial-grammar-based sequence as well as a repeating pattern. Most importantly, analysis of the data using secondary task error rate as a grouping criterion suggested that, in contrast to recent claims (Stader, 1995, Stader & Neely, 1997; Frensch, Buchner & Lin, 1994, Frensch, Lin & Buchner, 1997), attention appears to be a critical factor in sequence learning. This is the topic to be pursued further in this chapter. Two experiments, Experiments 2a and 2b, were carried out. Experiment 2a was identical in design and procedures to Experiment 1 except that the Rep and Gram sequences were used as each others transfer sequence. In Experiment 2b, a practice session was added to acquaint subjects with the requirements of the secondary task. The specifics will be noted below. As will be shown, this addition of the practice session influenced the way subjects allocated their attentional resources to the tasks, which had significant impact on sequence learning. In addition, a subtle change in the timing parameters of the secondary task was also made. As this factor

is one that will be addressed more extensively in another experiment, it will be discussed in greater detail later.

4.1 Experiment 2a

4.1.1 Methods

Subjects. Seventy-one subjects (29 males, 42 females) were recruited from Brooklyn College of CUNY. The students participated in the experiment to satisfy a course requirement. The subjects' age ranged from 17 to 45 ($M = 23.83$, $sd = 6.19$). Applying a set of screening criteria to eliminate subjects who were not following instructions (see details in Experiment 1a) resulted in an effective sample of 59 subjects (23 males, 36 females). It appears that more subjects were screened out in this experiment. One reason is that a larger sample was recruited.

Equipment and Material. The same computers were used to collect the data in this experiment as in previous experiment. The software was modified slightly such that the pseudo-random transfer sequence was eliminated. Instead, the Gram and the Rep sequences were used as each other's control sequence. In other words, if subjects were trained with the Gram sequence, they were transferred to the Rep sequence in Block 7, and vice versa for subjects trained with the Rep sequence. Otherwise, the experiment is a direct replication of Experiment 1.

One of the goals of this experiment is to further explore the nature of sequence learning. Subjects may simply pick up the relative frequencies of distinct events and respond accordingly. To rule out the possibility that sequence learning can be accounted for exclusively by sensitivity to simple event frequencies, one of two

Table 4.1 The statistical properties (simple event frequencies and 1st-order conditionals) of the (Rep and Gram) sequences used in Exps 2a and 2b.

Simple Event Frequencies in the Sequences:										
Rep "A" = 0.17, "B" = 0.17, "C" = 0.25, "D" = 0.25, "E" = 0.17										
Gram "A" = 0.16, "B" = 0.18, "C" = 0.21, "D" = 0.29, "E" = 0.17										
No. of Occurrence of 1st-order Conditionals:										
Rep					Gram					
X	BA	CA	DA	EA	X	BA	CA	DA	EA	
AB 10	X	CB 9	DB	EB	AB	X	CB	DB	EB	21
AC	BC 10	X	DC 10	EC 10	AC 8	BC 8	X	DC 3	EC 5	
AD 10	BD 10	CD 10	X	ED	AD 11	BD	CD 9	X	ED 15	
AE	BE	CE 10	DE 10	X	AE	BE	4	CE 6	DE 10	X
20	20	29	30	20	19	21	25	34	20	

manipulations could be attempted. One is to select a sequence in which all events occur with equal probability. As is shown in Table 4.1, this was not achieved. Some events (C and D) occurred more often than others (A, B and E). Another manipulation is to select training and transfer sequences that are matched in event frequencies. This was done as closely as possible in Exps 1a and 1b. Since learning was still observed, it must be due to other characteristics of the sequence beyond simple event frequencies. The next step is to look at more complex events such as first-order conditionals. Exps 1a and 1b had a transfer sequence that matched the training sequences (Rep and Gram) in terms of the frequency of each type of conditionals (e.g., A- type, which may include AB, AC, AD, and AE). However, the transfer sequence had more instances of each type than the training sequences. Note that the two training sequences (Rep and Gram), although not perfect, are better matched with respect to the number of instances of each type of conditionals. Under

these circumstances, learning, if observed, is attributable to higher order event contingencies. More convincing evidence will be presented in the next chapter where closer matching is achieved. In addition, the Rep and Gram sequences are differently structured. By pitting the Rep and Gram sequences against each other, we may reveal if one is easier to learn than the other. For instance, if Rep is better learned than Gram, then transferring from Rep to Gram will show a greater disruption than from Gram to Rep.

Design and Procedure. The same basic factorial design was used in this experiment as before. Three factors were included: namely, sequence type (SeqType), modality (Mod) and attention (Attent), which, as noted previously, is a subject factor based on breaking subjects into Good Counters and Poor Counters. Approximately half of the subjects were assigned randomly to each of the first two conditions. Other aspects of the experimental procedure were kept identical. Subjects were first trained in six blocks of 120 trials each. The sequence was changed in the 7th block and then changed back in the 8th. The experiment ended with a prediction block and a short questionnaire to probe subjects' conscious knowledge about certain aspects of the stimulus environment including the nature of the primary stimulus sequence.

4.1.2 Data Screening

The same data screening criteria used in Experiment 1a were applied here before analyses were performed. First, subjects whose average PC_{R1} across all trial blocks were under 95% were removed. In addition, subjects who had average RT during the training blocks greater than or equal to 1,000 ms were also removed. Outliers on the learning measure ($RT_7 - RT_6$) were scores outside the range defined by two cut-off

points. The lower bound is equal to the lower hinge - 1.5 IQR and upper bound is equal to upper hinge + 1.5 IQR (for a more detailed explanation, please see Chapter 3). In the original data, average PC_{R1} was 97.86% (sd = 1.87%, n = 71). After screening, it was 98.16% (sd = 1.28%). Screening resulted in eliminating 12 subjects (17%) from the original sample. Five of these subjects had average PC_{R2} s less than 95% and six had average RTs greater than or equal to 1,000 ms. But those failing the PC_{R1} all had scores greater than 90%. There was one subjects who showed a RT decrement from Block 6 to Block 7, ($RT_7 - RT_6$), of 520 ms. As a few "out of range" cases were identified in the next experiment (Exp. 2b), data records of each case were examined to gain some insight as to the cause of the "anomaly". The most likely cause, it seems, is due to a change of strategy or emphasis, rather than lapses of attention or fatigue, either in the block prior to the sequence change, i.e., Block 6, or in the transfer block, i.e., Block 7. To substantiate this statement, let me note that in all cases, the slower block was invariably accompanied by a higher accuracy either on the primary or secondary task, or both. For instance, an unusually large increase of RT in Block 6 was accompanied by greater accuracy on the secondary task suggesting that the subject temporarily shifted more attention to the distractor in that block. He or she may return to his or her old way of responding in the ensuing (transfer) block. This would result in a big decrement in RT from Block 6 to Block 7. In contrast, similar shift of emphasis may take place in the transfer block (Block 7) producing an unusually large RT in that block. Both of these situations yielded RT changes that either mask or inflate the real effect of learning. As the observed effect in terms of a decrement of RT is typically small within a certain range, those extreme scores due to

irrelevant and intractable causes grossly increase the variability within the group and distort the true magnitude of learning. It seems justifiable to remove those cases before data analysis was performed.

4.1.3 Results and Discussion

One goal of this experiment is to replicate the basic findings of Exps 1a and 1b that learning takes place under the conditions. Fig. 4.1 demonstrates subjects' performance with the RT data plotted in two separate graphs, one for each sequence type (SeqType). Data from the repeating pattern (Rep) is shown in the left panel and that from the grammatical sequence (Gram) is shown in the right panel. Each graph reflects the performance of subjects at each level of the two factors examined in this experiment, namely, the modality of the secondary stimulus (Mod), and how well subjects attended to the secondary task (Attent). The general pattern of performance as shown in the RT data in corresponding conditions appears to be very similar to what was observed in Exps 1a and 1b. For the sake of clarity, the results will be presented and discussed in several sections.

RT measures and implicit learning

In Chapter 3, data were analyzed and reported to show how the factors affected subjects' overall performance during training as well as their effect on learning. In this chapter, I will focus on the effect of these factors on sequence learning.

Therefore, analysis on the training part of the data will not be presented.

It is obvious from Fig. 4.1, the disruptive effect of sequence change in Block 7 is unmistakable, which strongly suggests that learning took place. A 4-way ANOVA with Attent, SeqType, and Mod as the between-subjects factors and Block (6-7) a

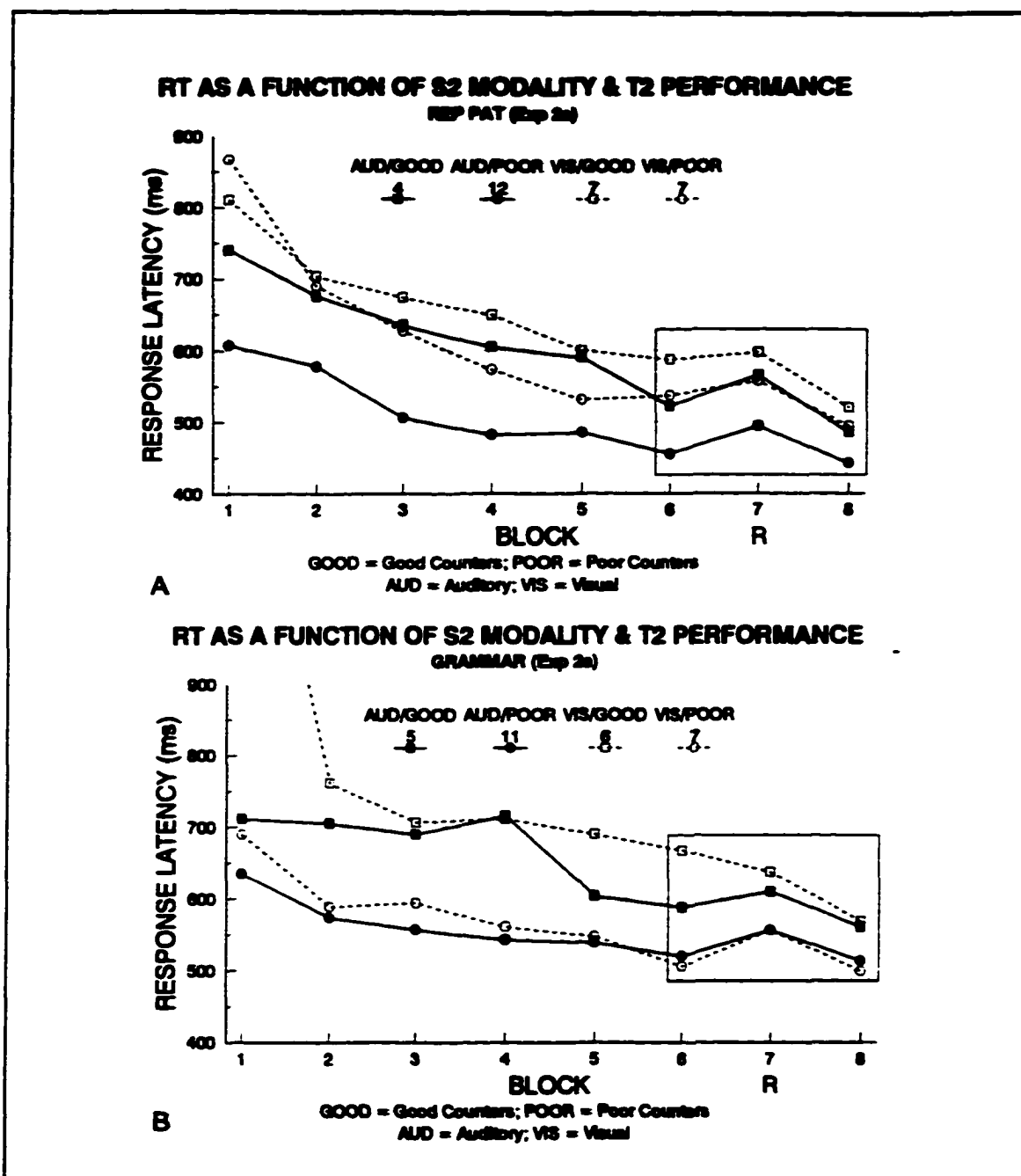


Figure 4.1 RTs are plotted by secondary stimulus modality and by how the secondary task was performed. Left panel: with a repeating pattern; Right panel: with a grammatical sequence.

repeated measure within-subjects factor was performed. The main effects of SeqType ($F_{(1,51)} = 1.31$, $MSe = 32771$, $p = .257$) and Mod ($F_{(1,51)} = 1.38$, $p = .245$) were not significant. But the main effects of Attent ($F_{(1,51)} = 4.38$, $p = .041$) and Block

$(F_{(1,51)} = 12.07, p = .001)$ reached significance. There was a marginally significant effect of the **Attent \times Block** interaction ($F_{(1,51)} = 1.31, MSe = 1226, p = .071$). None of the other interactions were significant ($F_{(1,51)} = 2.71, p = .105$ for **Mod \times Block**; all others, $F_s < 1.39, p_s > .244$). The **Block** main effect demonstrates a significant negative transfer effect due sequence change, which is a sign of learning. The marginal **Attent \times Block** effect suggests that the **Good Counters** did not learn as well as the **Poor Counters**. But the present experiment did not produce sufficient evidence to support this latter statement.

To see how these factors affected the performance of the primary and secondary tasks, the two corresponding measures, PC_{r1} and E_{r2} , were subject to two separate 4-way (**Attent \times Mod \times SeqType \times Block**) ANOVAs across all trial blocks. For the primary task accuracies, the analysis showed a significant main effect of **Attent** ($F_{(1,51)} = 6.60, MSe = 12.49, p = .013$) and of **Block** ($F_{(7,357)} = 2.17, MSe = 4.74, p = .036$) but not of **Mod** ($F_{(1,51)} = 1.96, p = .168$) and **SeqType** ($F_{(1,51)} = 1.10, p = .299$). The interactions were not significant except for **Mod \times Block** ($F_{(7,357)} = 2.74, p = .009$). The **Block** effect reflects some fluctuations over the blocks. The **Attent** effect indicates that **Good Counters** were generally more accurate on the primary task than **Poor Counters**. For the secondary task error rates, only the main effects of **Attent** ($F_{(1,51)} = 46.29, MSe = 3154, p < .001$) and **Block** ($F_{(7,357)} = 15.39, MSe = 178, p < .001$) were significant. No other effects reached significance. The difference due to **Attent** is not interesting here because it was created by grouping. The **Block** effect reflects the fact that accuracies were relatively lower in the two initial blocks

Pred as a measure of explicit knowledge

Explicit knowledge is assessed by percentage of correct prediction on the prediction task (Pred). In this experiment, subjects showed an accuracy of 27.46% (sd = 13.26), which is not significantly different from chance ($t_{(58)} = 1.41, p > .10$). In addition, the correlation between Pred and the implicit learning measure (i.e., $RT_7 - RT_6$) was not significant ($r = .08, p = .560, n = 59$). Again, as was found in Experiment 1, there is no evidence showing that subjects acquired explicit knowledge about the sequence, and even if some individuals did have some explicit knowledge, it was independent of their implicit knowledge. A 3-way (Attent \times Seqtype \times Mod) ANOVA showed a marginal effect of SeqType ($F_{(1,51)} = 4.00, MSe = 144, p = .051$), indicating that the repeating pattern produced slightly higher prediction accuracy. Other main effects were not significant ($F_s < 1, p > .485$). Neither were the interactions except that of Attent \times SeqType ($F_{(1,51)} = 10.01, p = .003$). The theoretical significance of this effect is not immediately obvious, however. Relevant is the result that Poor Counters (Pred = 28.22%, sd = 14.46%, n = 37) achieved no more explicit knowledge than Good Counters (Pred = 26.18%, sd = 11.14%, n = 22) (for Attent: $F_{(1,51)} = .49, p = .486$).

Section summary To recapitulate, Exp. 2a investigated effect of three factors on sequence learning through subjects performance on a SRT task with a concurrent secondary task. There is little doubt that subjects acquired some information about the sequence which facilitated their manual responses to the primary stimulus. Recall that in this experiment, the Rep and Gram sequences were used as each others transfer

sequence. As these two sequences were closely matched in terms of simple event frequencies, it can be said with greater confidence that improvement in task performance can not be accounted for by, and only by, noting differential frequencies amongst the events. As to the extent of learning in the two types of sequences, again, no evidence was found to substantiate the conjecture that the implicit learning mechanism is more efficient in encoding the repeating sequence than the sequence generated by an artificial grammar. The experiment also failed to reveal any difference in sequence learning due to the modality of the secondary task. Moreover, there is no evidence indicating that implicit learning assessed by a sequence transfer effect is mediated by explicit knowledge. However, the effect of attention on sequence learning did not receive sufficient support in this experiment. One possibility is that the manipulation of the availability of attention for sequence learning through the secondary task, mediated by the subject factor of how well they attended to the distractor task, in other words, the subjects might not have allocated sufficient attention to the secondary task for sequence learning to suffer. This hypothesis is explored in the next experiment.

4.2 Experiment 2b

To some extent, this experiment was intended as a pilot study for the experiment to be presented in the next chapter. Parenthetically, an exploratory data analysis revealed a slight but consistent difference in subjects' RTs sorted according to the onset timing of the secondary stimulus. It would be interesting to investigate the effect of the timing of the secondary stimulus relative to the onset of the primary stimulus

on sequence learning. However, although the timing was varied in previous experiments dating back to Cohen et al. (1990), it was not manipulated systematically. In order to observe the effect of the timing factor, it entails some kind of consistent onset scheduling. Following Cohen and his colleagues, I adopted the 40, 80 and 120 ms. onset relative to subject's response in the experiments reported thus far. Knowing that the effect, if it can be shown, is likely to be very small and subtle, we need to maximize the difference. To do so within the existing framework, the software must be modified in two ways. One change involved the parameter settings of the onset of the secondary stimulus. In Exps 1a and 1b, and in Exp. 2a, these were 40, 80 and 120 ms with a secondary stimulus duration of 80 ms. placed in an RSI interval of 200 ms. In this experiment, they were changed to 50, 100 and 150 ms. To keep the RSI unchanged, the duration of the secondary stimulus was reduced to 50 ms. These subtle changes in the procedure of the experiment, however, were not expected to affect subjects' performance in any foreseeable way. On the other hand, an apparently minor procedural change was introduced as a remedy to a problem emerged from previous experiments. Specifically, post-experimental interviews revealed that some subjects may have failed to understand, or follow, the instructions as to how the secondary task should be performed. Or they simply skipped through the instructions without trying to read them. One subject, for example, admitted to the failure to read the instructions carefully. She was confused when the computer prompted her to report how many times the target secondary stimulus occurred. Another subjects reported that she was not sure what the critical characteristics the target stimulus bore and counted a warning signal when a wrong key was pressed for the SRT task. To

minimize failures of these kinds,, a short practice session of 20 trials on the secondary stimulus counting task was added in this experiment. It was implemented in a conditional recursive algorithm in that it required subjects to produce a correct count before allowed to continue. For example, if the target signal was presented 12 times (determined by a random routine), the subject must enter an exact number to pass the test. Any other numerical input would trigger another practice session. This change resulted in a contrast to Exp. 2a, which constitutes another way to manipulate the availability of attention. This study holds the promise to illuminate the issue of attention.

4.2.1 Method

Subjects. Eighty subjects (16 males, 64 females) were recruited from the same subject pool but different students were sampled. They participated to fulfill the requirement of a course. The mean age of the subjects is 19.84 years ($sd = 4.85$) ranging from 15 to 55. After screening, a sample of 59 subjects (12 males, 47 females) was retained.

Equipment and Material The hardware used to run the experiment remained unchanged. The software was the same used in Exp. 2a except for the changes noted in the introduction of this experiment.

Design and Procedure The design features in this experiment remain unchanged from Exp. 2a. The procedural changes were discussed in the introduction section.

4.2.2 Data Screening

The same data screening criteria were applied in this experiment. Before screening, the average PC_{R1} was 98.01 % ($sd = 1.90\%$). After screening, it was

98.32% ($sd = 1.21\%$). Screening removed 21 subjects from the original sample. Six of them failed the PC_{R1} criterion although lowest score was 92.9%. The loss of subjects was not even with regard to the experimental conditions. The RT criterion resulted in rejecting 10 subjects. This attrition may not be independent of the manipulation of attention through the practice session. It is conceivable that the exactitude of the practice criterion may have suggested to the subjects that the same may be true for the rest of the experiment. This may have caused (at least some) subjects to slow down on the SRT task in order to attain greater accuracy on the target count. As it turned out, all 10 subjects would have been categorized as Good Counters. The remaining five subjects were removed because they contributed out-of-range values on the implicit learning measure ($RT_7 - RT_6$). This may or may not result from the increased perceived demand of the task. This distribution, however, did not seem to be biased one way or another.

4.2.3 Results and Discussion

RTs are plotted in Fig. 4.2 to demonstrate the effects of three factors, namely, SeqType, Mod, and Attent. Data associated with the Rep sequence is presented in the left panel while those associated with the Gram sequence is in the right panel. A 4-way ANOVA (Attent, SeqType, Mod and Block) for Blocks 6 and 7 produced significant main effects of Attent ($F_{(1,51)} = 7.95$, $MSe = 27079$, $p = .007$). No significant main effect was in SeqType ($F_{(1,51)} = 2.58$, $p = .115$), Mod ($F_{(1,51)} = 1.57$, $p = .216$) or Block ($F_{(1,51)} = 1.80$, $MSe = 1105$, $p < .186$). None of the interactions reached significance ($F_s < 2.32$, $p_s > .133$). The lack of significance in the Block main effect indicates that sequence learning was probably not achieved in

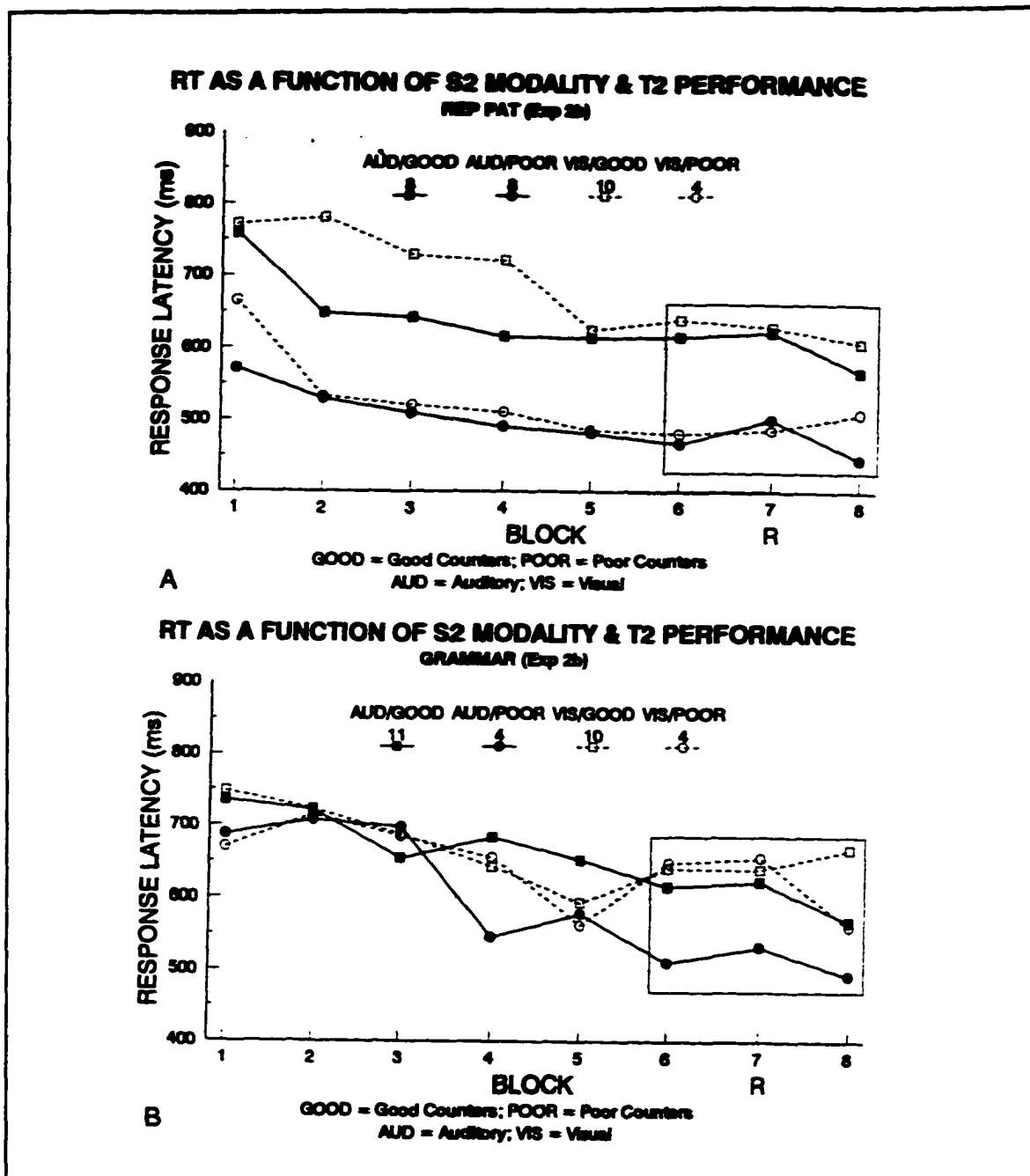


Figure 4.2 RTs are plotted by secondary stimulus modality and by how the secondary task was performed. Left panel: with a repeating pattern; Right panel: with a grammatical sequence.

this experiment. Although the curves of RT plotted in Fig. 4.2 exhibit characteristic increments in Block 7, which appears to be more pronounced in Poor Counters, the effect size, and/or the variability of the data warrant no claim of learning. To

conclude, adding a practice session for the secondary task reduced sequence learning substantially. Comparing this result with the finding in Exp. 2a suggests attention may be important in successful sequence learning.

Pred as a measure of explicit knowledge

In this experiment, subjects correctly predicted the stimulus 26.82% of the time ($sd = 13.15$), which is not reliably different from chance ($t_{(38)} = 1.05, p > .20$) given a chance probability of .25. Correlation between implicit ($RT_7 - RT_6$) and explicit (Pred) learning measures showed that the two were independent of one another ($r = .14, p = .281$). The Poor Counters (Pred = 29.26, $sd = 13.19, n = 20$) enjoyed a slight numerical advantage over the Poor Counters (Pred = 25.57, $sd = 13.13, n = 39$). A 3-way ANOVA with SeqType, Mod and Attent as the factors produced no significant main effects ($F_s < 1, p_s > .356$) nor interactions ($F_s < 1.72, p_s > .196$) This is consistent with the result obtained in Exp 2a.

4.3 General Discussion

At this point, a comparison between Experiment 2a with Experiment 2b would be helpful to appreciate the importance of attention in sequence learning. To evaluate the role of secondary task performance in sequence learning, let us take a closer look at how the secondary task practice affected subjects' policy of attention allocation. The error rates were plotted across the trial blocks in Fig. 4.3 for subjects in Exps 2a and 2b. It clearly shows that subjects' counts of the target secondary task signal, on average, were much closer to the actual numbers when given the practice task. To look at the data in a different way, Fig.4.4a and b show the distributions of the

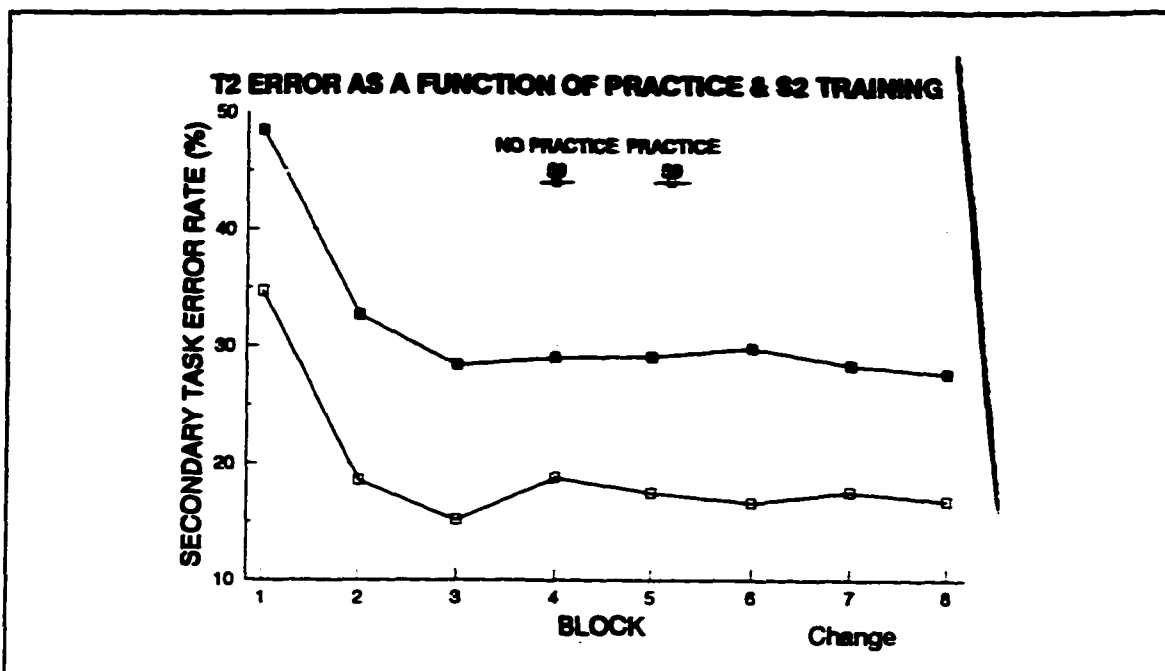


Figure 4.3 Plotted are mean secondary task error rates (E_{R2}) across the trial blocks in Experiments 2a (filled squares) and 2b (empty squares).

secondary task error rate (E_{R2}) averaged across the Blocks 3-8 in Exps 2a and 2b respectively. A comparison between the two charts makes it clear that requiring subjects to practice on the secondary task before data collection indeed produced superior performance on the task. But, as will be shown later, this higher level of accuracy was attained at a considerable expense of the processing of the primary task. Of special interest is that mobilizing attention to the secondary task made it less available to the processes in sequence encoding.

Two strategies were employed in this experiment to explore the effect of attention on sequence learning. As in Exps 1a and 1b, subjects were divided into two groups according to how well they performed the secondary task. This reflects the way subjects allotted attention to the secondary task. I'll refer to attention available for sequence encoding influenced by this factor *endogenously determined*. On the other

hand, specific to this experiment is the manipulation of attention by having, or not having, subjects practice on the secondary counting task. It is assumed that this practice biased subjects to allocate more attention to the secondary task. I will call attention available determined by the presence or absence of an external factor (e.g., special instructions or environmental constraint) *exogenously induced*. The Good Counters vs. Poor Counters contrast represents an endogenous attentional factor while the contrast between Exps 2a and 2b represents an exogenous attentional factor. To

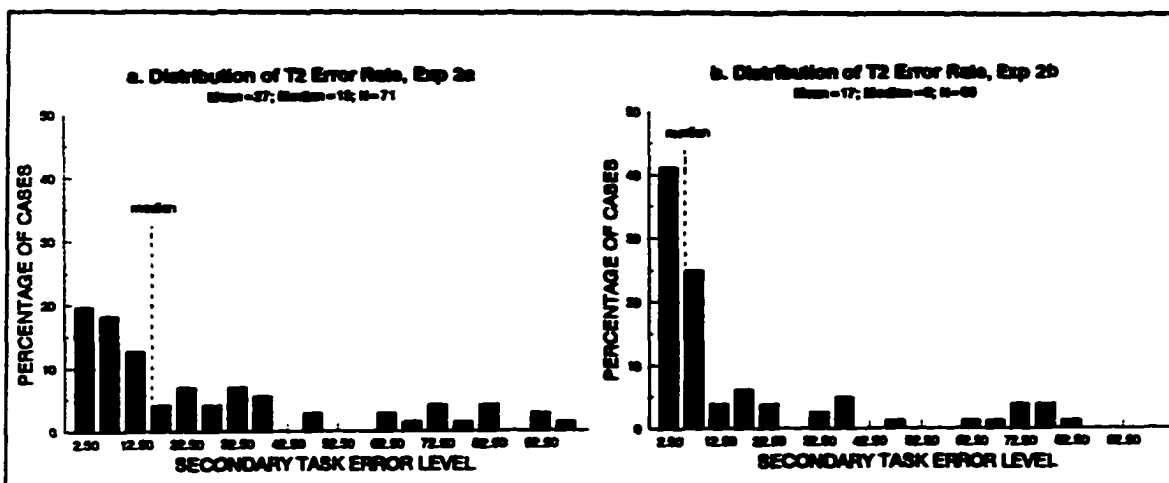


Figure 4.4 a. Distribution of averaged secondary task error rate in Experiment 2a; b. Distribution of averaged secondary task error rate in Experiment 2b.

facilitate the understanding of the relationship between these two attentional factors, RTs from Exps 2a and 2b are plotted in Fig. 4.5 along two dimensions. The endogenous attentional factor is marked by the divide between the square symbols (Good Counters) and circular symbols (Poor Counters). The exogenous dimension is distinguished by filled and empty symbols. Overall, Poor Counters responded consistently faster than Good Counters ($F_{(1,114)} = 21.58$, $MSe = 130902$, $p < .001$). They also demonstrated better learning ($F_{(1,114)} = 7.10$, $MSe = 1142$, $p = .009$).

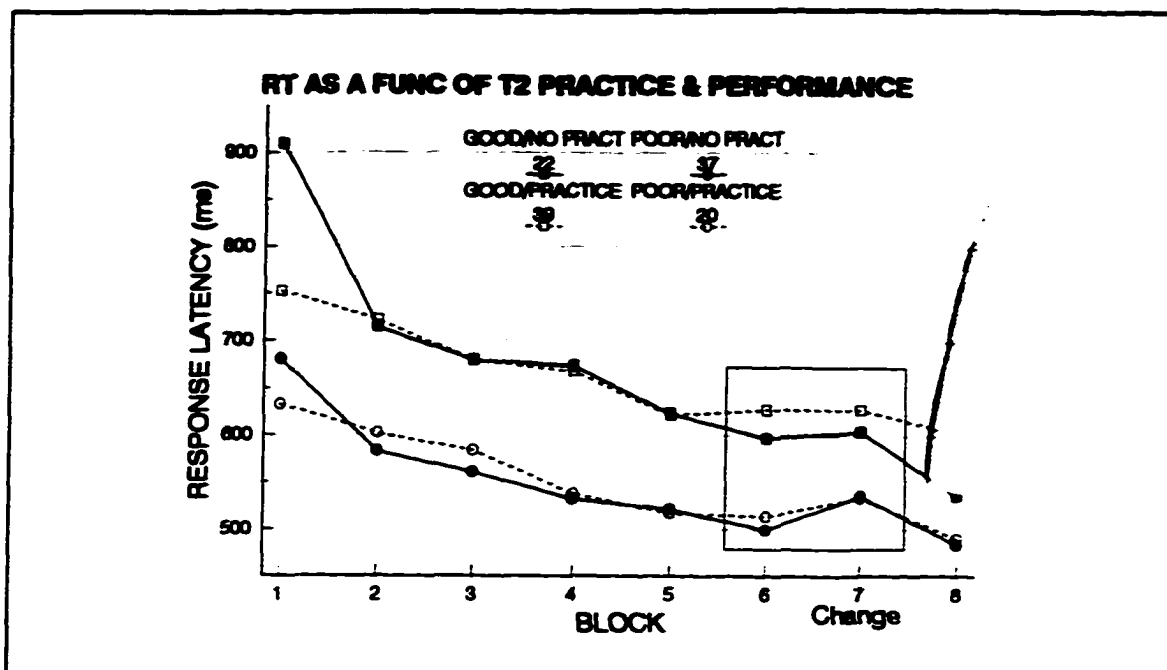


Figure 4.5 RTs for Good Counters and Poor Counters are plotted across the trial blocks for data from Exp 2a (without secondary task practice) and Exp 2b (with practice).

Note, however, that subjects in the two experiments displayed a very similar pattern of performance except in Block 1. In the initial block, subjects who practiced the secondary task exhibited faster RTs on the primary task, indicating that mastery of the secondary task facilitated the performance of the primary task ($F_{(1,114)} = 5.55$, $MSe = 32606$, $p < .020$). However, it did not translate into better learning ($F_{(1,114)} = 1.63$, $MSe = 1142$, $p = .25$). On the contrary, practice appeared to have impaired learning. These observations suggest that subjects' behavior is affected by how attentional resources are allocated in a complicated manner, either by internal mechanisms or external influence (such as a [perceived] emphasis on task demand). The effect of attention on sequence learning per se is revealed more lucidly by Fig. 4.6. In this bar chart, the height of a bar is determined by an $RT_7 - RT_6$ as shown in Fig. 4.5. It is more obvious here that Poor Counters achieved far better learning than

Good Counters. This is consistent with what was found in Experiment 1. The figure also suggests that subjects who did not receive practice on the secondary task produced better learning than subjects who did, but this was not supported statistically ($F_{(1,114)} = 1.63, p = .205$).

Reversal-Trio Analysis and Inhibition of Return One issue that is germane to the present study and also of interest to other psychologists is how the structure of the

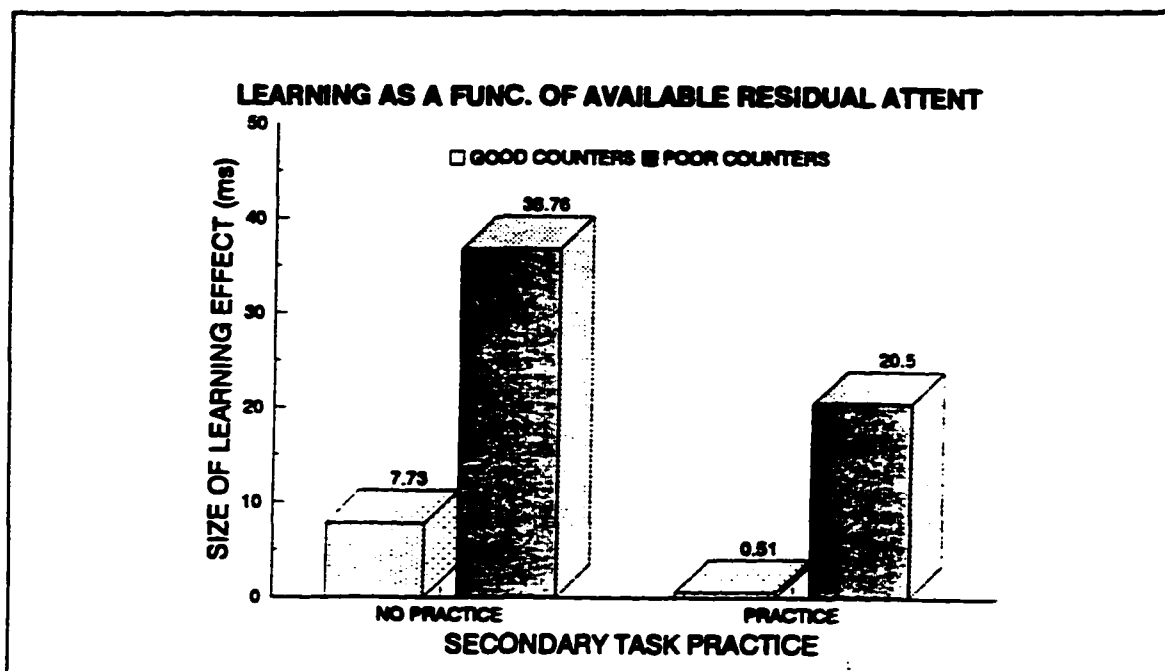


Figure 4.6 Effect of learning as a function of availability of attention controlled by internal and external conditions.

event sequence affect subjects' behavior. Certain peculiar microstructures in the sequence may affect subjects' response in predictable ways. One such structure is comprised of event chains involving a repetition of an event separated by a different event (e.g., CEC, or ADA). Some researchers call them "returns", but I prefer to call them reversal-trios in this study. It was shown previously that subjects were slower in their response to the returning event, or the second leg of the reversal trio

as compared to non-reversals (Tipper, Weaver & Watson, 1996). This phenomenon has been referred to as the *inhibition of return*. This effect of the reversals was documented in this experiment as well. In general, subjects took more time to respond on trials where the reversal fell compared to the non-reversal trials. What is interesting is that attention seems to play a role in this phenomenon.

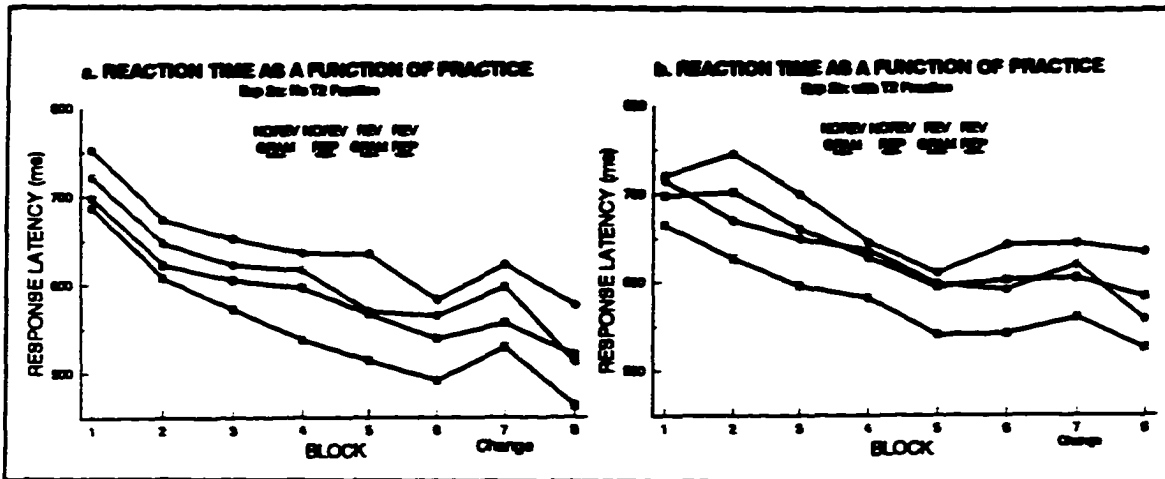


Figure 4.7 Plotted are RTs in Exp 2a (Panel a) and 2b (Panel b) by type of trio structure defined by whether it contains a reversal.

Data were analyzed using a 2 (Attent: Good Counters vs. Poor Counters) x 2 (structure: Reversal vs. Nonreversal) x 8 (Block) ANOVA. As the proportion of reversal trials was different in the GRAM (1/5) sequence and REP (1/4) sequence, data must be analyzed separately. Apart from that, the two experiments were also analyzed separately. This resulted in four ANOVAs. For Experiment 2a, the Attent main effects are (or at least, marginally) significant for both Gram and Rep: $F_{(1,27)} = 5.01$, $MSe = 309581$, $p = .034$ and $F_{(1,27)} = 4.19$, $MSe = 205622$, $p = .050$ respectively). The Structure main effects are also significant: $F_{(1,27)} = 36.97$, $MSe = 5552$, $p < .001$ and $F_{(1,27)} = 63.77$, $MSe = 3772$, $p < .001$) and so are the interactions between Block and Structure: $F_{(7,189)} = 2.98$, $MSe = 3757$, $p = .005$

and $F_{(7, 196)} = 4.82$, $MSe = 3586$, $p < .001$). The interactions reflect the fact that the difference in Reversals and Nonreversals grew larger as training proceeded (see Fig. 4.7 a and b).

The picture in Experiment 2b is somewhat different. First, the **Attent** main effect was significant with the **Rep** sequence, $F_{(1,28)} = 11.35$, $MSe = 222461$, $p = .002$. Second, the **Structure** main effect was also only significant in **Rep**, $F_{(1,28)} = 9.53$, $MSe = 8690$, $p = .005$). Third, the **Block by Structure** interaction was not significant ($F_{(7, 196)} = 1.50$, $MSe = 4458$, $p = .170$) (see Fig. 4.6 c and d).

The absence of a difference between **Good Counters** and **Poor Counters** in **Gram** in Experiment 2b and a smaller one in **Gram** in Experiment 2a suggest that the inhibition is tied to the nature of the sequence. The specific information that subjects may have learned is that the occurrence of reversal trios are much less frequent than non-reversals in this experiment. As they learned more about this with practice, they became less ready to respond to a reversal when it did show up. It should be noted that knowledge involving event trios rely on some appreciation of second-order contingencies. Subjects must know (implicitly) that when a event occurred, and it was followed by another event, that first event is less likely to occur in the next trial. As the data show, acquiring this kind of information may require attention.

4.4 Chapter Summary and Conclusion.

Experiments 2a and 2b confirmed that a sequence made up entirely of non-unique transitions could be learned implicitly even when the SRT task was performed together with a distractor. Such learning goes beyond simple event frequencies, and is

probably more complex than first-order contingencies. Apart from that, sequence learning seemed to be a function of how well subjects attend to the secondary task. However, attempts to manipulating attention by external means initially indicated that giving subjects a practice on the secondary task resulted in poorer sequence learning. But analysis performed on combined data failed to show a Practice-No Practice effect on learning. It only seemed to affect subjects' response efficiency at the beginning of the training phase. Its effect on sequence learning was not substantiated in this experiment. Concerning the implicit-explicit nature of the sequence knowledge, the present study failed to support the explicit stance. It is likely that some subjects may have achieved some explicit knowledge of the sequences to which they were exposed extensively. However, insofar as the implicit and the explicit measures do not tap on a common database, the presence of explicit knowledge does not undermine the statement that sequence learning is achieved primarily implicitly.

As noted in the introduction to this experiment, an exploratory analysis on the RT data by sorting out RTs according to which stimulus onsets they were associated with showed that subjects' response rate on the SRT task was differentially affected by the onset of the secondary stimulus. This observation is consistent with the general Bottleneck theory. It indicates that at least a certain processing stage of each task must be handled by a central mechanism. The consequence is that when one task is being processed, the other has to wait (Pashler, 1993). Unfortunately, the analysis has little to offer as to how the different onsets might have affected sequence learning per se since the onset was determined randomly. A systematic manipulation of the onset of the secondary stimulus relative to that of the primary stimulus is called for in order to

tease apart the specific effect on sequence learning. At this point, it can only be speculated that the closer the two tasks were queued for processing, the less likely attention will be available for encoding the relationship between adjacent events. This is the topic for the next chapter.

CHAPTER 5

EXPERIMENT 3: SEQUENCE LEARNING AND ATTENTION SHARING

In Experiment 1, subjects who were Good Counters demonstrated poorer sequence learning than those who were Poor Counters, suggesting that the Good Counters, for one reason or another, allocated more attention to the secondary task. However, the Poor Counters, again for one reason or another, chose, or simply happened, to give less attention to that task. Either interpretation fits with the assumption that, for any individual, attention has a unit capacity (Pashler, 1994). As it turned out, the manner in which subjects allocated their attention to each of the competing tasks resulted in performance of these tasks with qualities directly reflecting their biases. Experiment 2 showed that subjects who received preliminary training on the secondary task achieved less sequence learning than those who didn't. The suggestion is that the practice induced subjects to allocate more attention to the distractor task leaving less attentional resources for encoding the structural information underlying the sequence of primary task events. Whether attention allocation was endogenously initiated or exogenously induced, the result is that sequence learning suffers. This clearly suggests that attention plays a critical role in the situation.

However, there are two methodological weaknesses in the first two sets of experiments. Specifically, the critical variable in Exps 1a and 1b involved a subject factor and the Practice/No Practice distinction was observed across experiments rather

than manipulated systematically within the same experiment. As such, both limit our ability to infer unequivocally a causal relationship between attention and the amount of learning. Recall that in Experiments 1 and 2, it was argued that subjects having high error rates on the secondary task were not paying sufficient attention to that task thus leaving more attention for sequence encoding. It was suggested these subjects, instead of keeping track of the occurrence of the target secondary stimulus on a trial-to-trial basis as they were instructed to, were simply making rough estimates at the end of each block of trials. A simple way to examine this is to create a control condition in which subjects are specifically instructed to do just that, that is, to estimate.

In this experiment, the Good Counters/Poor Counters distinction was retained and an Estimators group was added. The Estimators were only asked to estimate the number of times the target signal occurred in a block of trials at the end of that block. No practice was given on the counting task for any subjects. Since estimation is considered to be much less demanding and exacting, subjects in this condition should be similar to the Poor Counters.

In the two preceding sets of experiments, as in virtually all sequence learning experiments using the SRT task with a distractor, the secondary task was presented using a fixed interval (usually, 200 ms) between subjects' manual response and the onset of the next primary stimulus. As noted earlier, researchers typically varied the onset of the secondary stimulus relative to subjects' key-pressing response just made. Since the response-stimulus interval of the SRT task is fixed, this means that the interval between the onset of the secondary stimulus and that of the primary stimulus

on the next trial was varying from trial to trial in uncontrolled ways. In this experiment, this factor will be called **SSPSOA**, which stands for **Secondary Stimulus-Primary Stimulus Onset Asynchrony**¹⁵. Focusing on the SSPSOA allows us to investigate the relationship between the processes of the secondary task and those of the primary task in that brief time frame.

The reason for taking this approach is simply that it is likely that not all stages of a cognitive process require the attention of the central executive mechanism. Pashler (1993) has shown that while stimulus perception and response execution can be carried out with little interference to each other, a stage between these two, namely, response selection, seems to require central attention. Pashler attributed the psychological refractory period (PRP) in the context of dual task performance to the bottle-neck in attention. His theory explains why the addition of the secondary task always slows down subjects' response to the primary task. But he did not explore the cognitive consequence of queuing the attention-demanding stage of two tasks in close succession. It is likely that the prolonged engagement of the working memory by the attentional mechanism precludes sequence learning.

This experiment makes use of the standard procedure in which the secondary stimulus is presented at different delay intervals following subjects' response to the primary stimulus. No one, however, has ever asked if different SSPSOA might affect

¹⁵ In other experiments on sequence learning such as Cohen et al. (1990), the onset time of the secondary stimulus was also varied. It was usually measured from the subjects' key-pressing response to the onset of the secondary stimulus. Since our focus is on the effect of the onset timing of S_2 on subjects' response efficiency to the next S_1 , we use the onset of S_1 on trial $t + 1$ rather than the onset of R_1 on trial t as the reference point. The timing settings, however, are equivalent to those in previous experiments.

sequence learning differently. One reason why the question has not been addressed may be due to the fact that the varied onsets were determined randomly. In this experiment, SSPSOA is manipulated as a between-subjects factor to see how this would affect sequence learning.

5.1 Method

Subjects A sample of 198 subjects (66 males and 132 females) was obtained from the same subject pool. The subjects' mean age was 21.93 ($sd = 6.56$) ranging from 16 to 49.

Equipment and Material The same computers were used in stimulus presentation and data collection. The computer program was modified to accommodate the changes in the procedures. In addition, new training and transfer sequences were created to provide better control of the statistical constraints of the sequences (see Table 5.1). The training sequence is made up of a randomly generated 12-element pattern recycled 10 times. The transfer sequence is a continuously generated 120-element pseudo-random sequence with properties closely matching those of the training sequence. These sequences possess statistical properties as shown in the table. Specifically, simple event frequencies were identical in both sequences. Although exact match could not be found, the two sequences were very similar in the kind and number of occurrence of first-order conditionals. As in Experiments 1 and 2, events were not allowed to repeat in two consecutive trials. In addition, the sequences were generated with another constraint in that reversal trios, e.g., ABA, or ECE were not allowed. Analysis in Experiment 2 showed that subjects tended to be slower on trials that formed the second leg of the reversal trio. This effect may be due to *inhibition of*

Table 5.1 Training and transfer sequences used in Experiment 3 and two properties of these sequences: simple event frequencies and 1st-order conditionals.

Training Sequence									
CEADCBDECDAB × 10 times									
Transfer Sequence									
CEADCBDCCEADC									
BDECBDABCEAD									
EADCBDEADCBD									
CEADCEADCBD									
BCDABCDECBD									
CDABCDABDCBD									
ABCEABCDABCE									
ADCEABCDECDE									
CDECDECDCDABCD									
ECDABDCEABCE									
Simple Event Frequencies in Both Sequences									
"A" = 0.17, "B" = 0.17, "C" = 0.25, "D" = 0.25, "E" = 0.17									
No. of Occurrence of 1st-order Conditionals (Training/Transfer):									
X		BA		CA		DA	10/ 9	EA	10/11
AB	10/12	X		CB	10/ 8	DB		EB	
AC		BC	9/10	X		DC	10/11	EC	10/ 8
AD	10/ 8	BD	10/10	CD	10/12	X		ED	
AE		BE		CE	10/10	DE	10/10	X	

return (see Tipper & Cranston, 1985) or may simply reflect the fact that such reversals occurred relatively infrequently. In either case, it seemed to make sense to remove a structural characteristic that is known to have a dampening effect and also interacts with attention in a complex way.

Design and Procedures In order to concentrate on the issue of attention, and to keep the complexities of the experiment manageable, the sequence type and modality factors were dropped in this experiment. Only the Repeating Pattern was used and all

subjects were instructed to attend to the auditory secondary signal. The experiment implemented a 3 (Attent: Good Counters vs. Poor Counters vs. Estimators) x 3 (SSPSOA: 150 vs. 100 vs 50 ms) × 10 (Block) mixed design with Attent and SSPSOA as between-subjects factors and trial blocks as a repeated measure within-subjects factor.

Several minor changes in the procedures from Experiments 1 and 2 were made. First, the training phase was extended from 6 blocks to 8 blocks to give subjects more exposure to the event sequence. Data from previous experiments indicated that subjects' RTs did not reach asymptote showing that learning was not completed by the end of the training phase. Second, the paper-and-pencil version of the questionnaire was replaced by appropriate questions integrated into the computer program. Subjects' responses were then coded and recorded to facilitate data analysis. For example, subjects were asked, immediately after the testing phase at the end of Block 10, whether they noticed any regularity in the primary (SRT) task event sequence. This change was intended to probe subjects' conscious awareness of the structured nature of the event sequence. They were also asked if they found the experiment interesting as an index of motivation. This represents an improvement over the post-experimental questionnaire because relevant information is elicited right after the experience and before it gets contaminated by other experiences. Third, a confidence rating with a scale between 1 and 7 was added after the subject entered his or her count of the secondary task at the end of each block. Fourth, in addition to the (cued) prediction task, a production task was added in which subjects were asked to produce the whole sequence by pressing the keys one at a time. This was intended as another (possibly

stronger) test of explicit knowledge.

5.2 Data Screening

The data were subject to the same screening procedures as in the earlier experiments. The average RT criterion removed three subjects. It turned out these were older subjects (age > 43). The average PC_{R1} criterion is responsible for the exclusion of 14 subjects. Reviewing the data showed that these were mostly Poor Counters but with learning scores ($RT_7 - RT_6$) greater than 0. In addition, Estimators whose accuracy on the secondary task fell within the range of Good Counters were eliminated. This group was intended as a control group to the Good Counters and Poor Counters collapsed by secondary task error rate (E_{R2}). However, 6 subjects had E_{R2} s below 10% and it was suspected that they might have actually counted the target signals. In view of this, these subjects were removed from the sample before analyses were performed¹⁶.

After data screening using all relevant criteria, the final sample included 170 subjects (53 males and 117 females) with a mean age of 21.74 (sd = 6.04). For the entire sample prior to data screening, average PC_{R1} was 97.73% (sd = 1.86%). Screening resulted in a sample with $PC_{R1} = 98.09\%$ (sd = 1.27%). For the non-estimators, this is 98.03% (sd = 1.29%) whereas for the Estimators, this is 98.25% (sd = 1.22%).

5.3 Results and Discussion

¹⁶ Analysis with the "Good Counters" in the Estimators group included was also performed. The results were not different for interpretation purpose. The only notable difference is it revealed a marginally significant main effect of SSPSOA ($F(2,166) = 2.78$, $MSe = 22732$, $p = .065$) in the learning effect analysis (Blocks 6-7).

RT measures and Implicit Learning

RTs are plotted in Figures 5.1 to show subjects' performance across the 10 trial blocks. Each graph represents how subjects performed the secondary task as a means of attentional "manipulation", namely, Good Counters, Poor Counters and Estimators. The three levels of the secondary stimulus onset (SSPSOA) were shown on each graph. A 3-way ANOVA, with Attent and SSPSOA as between-subjects factors and Block (1-8) as a within-subjects factor, was performed over the training blocks. The analysis yielded a significant Attent main effect ($F_{(2,161)} = 16.23$, $MSe = 99508$, $p < .001$), a significant SSPSOA main effect ($F_{(2,161)} = 3.11$, $p = .047$) and a significant Block main effect ($F_{(7,1127)} = 90.48$, $MSe = 3629$, $p < .001$). There was also a significant 3-way interaction (Attent \times SSPSOA \times Block: $F_{(36,1449)} = 1.86$, $MSe = 3376.84$, $p = .002$). None of the other interactions was significant ($F_s < 1.10$, $p > .359$). These results basically show that subjects performed differently at different level of each variable.

A 3-way (Attent \times SSPSOA \times Block) ANOVA involving Blocks 8 and 9 was run to examine to examine the specific impact of the Attent and SSPSOA variables on sequence learning. The analysis produced a significant Attent main effect ($F_{(2,161)} = 13.11$, $MSe = 22178$, $p < .001$), and a Block main effect ($F_{(1,161)} = 164.02$, $MSe = 1072$, $p < .001$). The main effect of SSPSOA did not reached significance ($F_{(2,161)} = 1.50$, $p = .226$). The interaction between Attent and Block, however, was marginally significant ($F_{(2,161)} = 2.55$, $p = .082$). Other effects were not significant. The Block effect indicates that learning was achieved overall, and, as is shown in Fig. 5.1, subjects who paid less attention to the secondary task (the Poor Counters and the

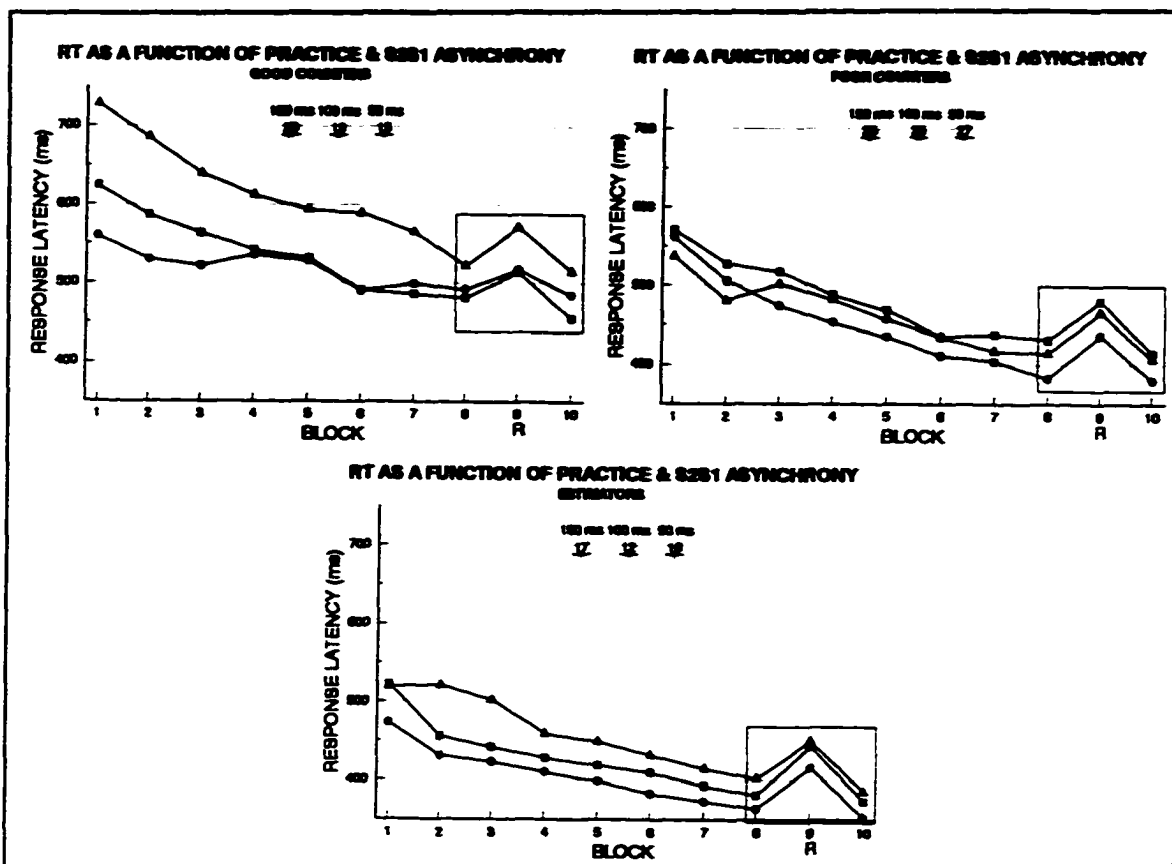


Figure 5.1 RTs are plotted across the trial blocks in a set of three graphs by Good Counters, Poor Counters, and Estimators respectively.

Estimators) exhibited better learning. However, the effect of attention on learning was not firmly supported by the data. In addition, the lack of interaction among SPSOA, Attent and Block affords not support to the prediction that a later S₂ or earlier onset would lead to compromised learning when subjects were severely distracted.

Admittedly, the failure to find the expected effect of stimulus onset on learning is disappointing, but the failure to find the previously observed effect of attention in this experiment is certainly bothersome. A little digging and fiddling were attempted to reconcile the discrepancy in findings between the present experiment and those conducted previously. Fig. 5.2 shows the size of learning effect broken down by each of the three levels of the Attent and SPSOA variables. This figure seems to give

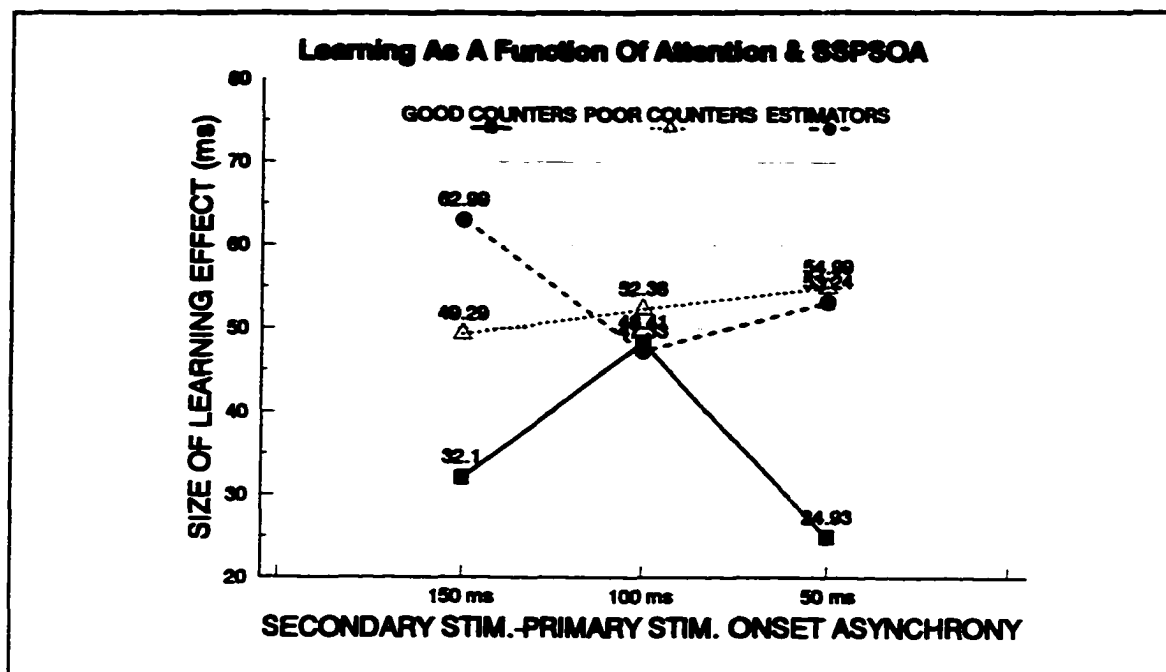


Figure 5.2 Sequence learning effects are plotted by allocation of attention and the onset of the secondary stimulus.

some hints on the subtle relationship between attention and the scheduling of the onset of the secondary stimulus with respect to sequence learning. As can be seen from Fig. 5.2, substantial learning was achieved by the Poor Counters and Estimators regardless of the onset of the secondary stimulus. The Good Counters, however, attained learning at a comparable level only when SSPSOA = 100 ms. Their learning was compromised when the secondary stimulus was scheduled either right after subjects responded to the primary task on the current trial (t) (SSPSOA = 150 ms) or immediately before the onset of the primary stimulus on the next trial ($t + 1$) (SSPSOA = 50 ms). Learning was especially subdued when the primary stimulus was immediately preceded by the distractor. In fact, this is the only condition in which the effect of learning is not reliably different from 0 (two-tailed $t_{(14)} = 1.73, p > .10$). This, of course, is the situation when attention is least available to sequence learning.

Learning was apparently not affected by when the secondary stimulus was presented when subjects were not following it closely. It is likely that the Poor Counters did not ignore the secondary task entirely. They probably processed the distractor stimulus perceptually. But they, as the Estimators, did not actually try to keep track of how many times the target signal occurred. This allowed them to have sufficient processing capacity for sequence encoding. The Good Counters, in contrast, tried to follow the instructions faithfully and counted the target signal to the best of their ability. By so doing, they dramatically reduced the amount of processing resources available for sequence learning. However, when the secondary stimulus is presented in such a way to allow attention to be shifted briefly to sequence encoding when a copy of the previous event is still available and the current event is identified, learning can still be achieved adequately. But if the secondary stimulus is presented soon after subjects made the manual response to the primary stimulus (SSPSOA = 150 ms), the attentional demand of the former forces processing resources to be channelled to the new stimulus. On the other hand, if the secondary stimulus is presented immediately before the primary stimulus (SSPSOA = 50 ms), both tasks compete for attention at the same time resulting in switching attention first to the secondary task and then to the primary task, and then back to the secondary task for updating the target count, which falls in the interval between the key-pressing response and the onset of the secondary stimulus. In both situations, sequence learning suffers severe impairment.

The above theorizing is irrelevant if the effect does not exist, as the results had shown. An obvious excuse is the delicateness of the observed effect and the lack of

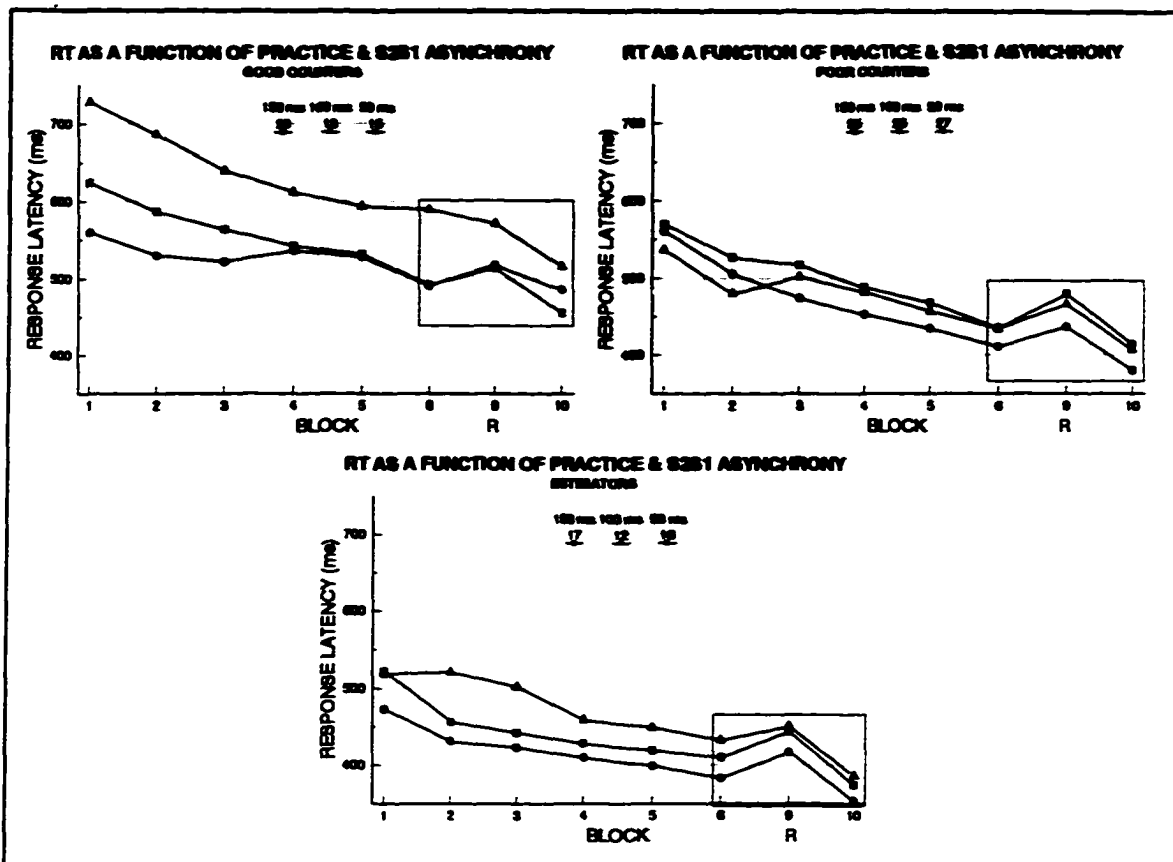


Figure 5.3 Figure reproduced based on data shown in Fig 5.1 with the data in Blocks 7 and 8 removed. The original Blocks 9 and 10 are now labeled 7 and 8.

power of the test. But part of the experiment concerns some pretty robust phenomenon documented relatively firmly in previous work. We had seen that generally Good Counters make poor learners. However, this is only marginally supported in this experiment. Did we see ghost there or did some thing go wrong here? To understand why the observed attentional effect on sequence learning was not observed in this experiment, an exploratory analysis using a 3-way (Attent \times SSPSOA \times Block: 6 and 9) ANOVA was performed on RTs of Blocks 6 and 9. The data shown in Fig. 5.1 were replotted in Fig. 5.3 with Blocks 7 and 8 removed and replace by the original Blocks 9 and 10. Recall that in the experiments reported in the two preceding chapters, subjects received six blocks of training on the SRT task

before learning was assessed. It is likely that learning was not developing at the same rate depending on the availability of attention.. However, after extensive training, subjects who had more attention reached asymptote in sequence learning while those who were learning slowly gradually caught up. Therefore, they show no difference in learning. If we imagine that training stopped at the end of Block 6 and learning was assessed in Block 7 rather than in Block 9. The test would be more like what we did in previous experiment. This quasi-analysis produced significant main effects of Attent ($F_{(1,161)} = 15.35$, $MSe = 19445$, $p < .001$) and Block ($F_{(1,161)} = 35.78$, $MSe = 1318$, $p < .001$) and a marginally significant SSPSOA effect ($F_{(1,161)} = 2.92$, $p < .057$). The 2-way interaction of Attent \times Block was now significant ($F_{(1,161)} = 15.35$, $MSe = 19445$, $p < .001$). The interaction of SSPSOA \times Block ($F_{(1,161)} = 15.35$, $MSe = 19445$, $p < .001$) was also marginally significant. Other effects were not significant. However, interpretation of the results was impeded because there is some difficulty in one aspect of the data. As can be seen from Fig. 5.3, subjects with the intermediate (100 ms) SSPSOA consistently responded more slowly than subjects with a longer (150 ms) or shorter (50 ms) SSPSOA. This observation appears to be counterintuitive and thus rather puzzling. One possible explanation may be due to individual differences. Since this factor is manipulated between-subjects, it is likely that more relatively slower subjects were assigned to this group than to the other groups.

Pred and Prod: two measures of explicit knowledge

Two procedures were used in this experiment to probe subjects' explicit sequence knowledge. One is the Prediction task in which subjects were prompted, on a portion

of the trials, to predict which stimulus was going to occur on the next trial. The other is the Production task in which subjects were required to produce the entire training sequence. The prediction task has been used widely and in the two sets of experiments reported in the previous chapters. The production task is new and is used for exploratory purpose only. For both tasks, percent correct ((number of correct responses / number of trials requiring a response) \times 100) was used as the performance index. Certainly there are problems concerning scoring for both tasks, but the difficulty is even greater for the production task. But those are empirical issues which must be pursued more seriously. For now, the two tasks yielded two measures, named Pred and Prod respectively. For the whole sample, mean Pred = 38.55% (sd = 16.98%) and mean Prod = 20.93% (sd = 5.33%). Given five alternative choices but only four was possible due the constraint that the same event could not occur on two consecutive trials, the probability of predicting/producing a correct event on any trial would be 25%. As it turned out, Pred was significantly better than chance ($t_{(169)} = 10.37, p < .001$) but Prod was below chance ($t_{(169)} = -12.02, p < .001$). The result based on Pred indicates that subjects may possess some explicit knowledge about the sequence. Prod, however, does not support this proposition.

One may question the validity of the prediction or production tasks as an instrument for measuring explicit knowledge. But an adequate discussion of this question is beyond the scope of the present study. However, some light can be shed on the issue by showing the relationship between these two explicit learning measures and the implicit measure, i.e., $RT_7 - RT_6$. The data were plotted, by comparison, in

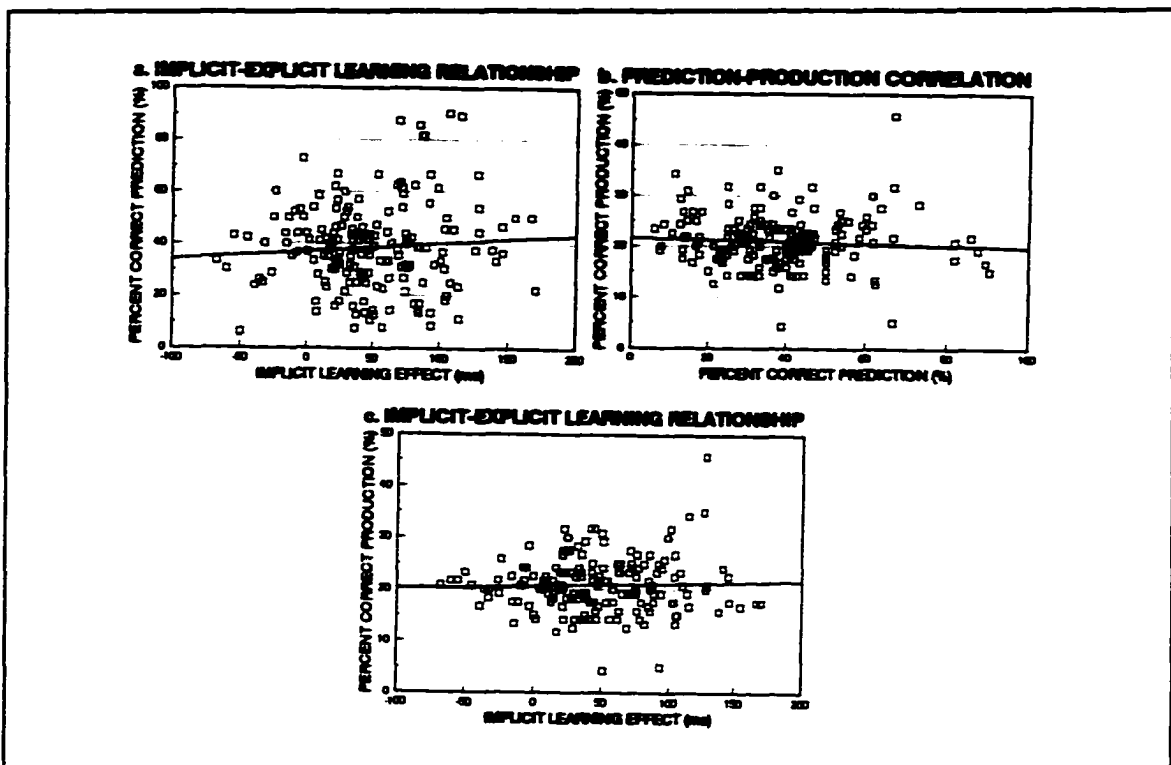


Figure 5.4 The implicit learning scores (RT_9-RT_0) and the two explicit knowledge scores (Pred and Prod) are plotted in three graphs to show the relationship between each pair. three graphs presented in Fig. 5.4. Pearson correlation coefficients were computed for each pair of the three measures: for Pred and the implicit learning, $r = .08$, $p = .272$ for Prod and (RT_9-RT_0), $r = .04$, $p = .609$ and for Pred and Prod, $r = -.06$, $r = .46$. None of these correlations is significantly different from 0. If we accept Pred and/or Prod as measures of explicit knowledge, once again, it is shown that subjects who do well on the implicit task may not necessarily do well on the explicit task. The independence of the two type of measures refutes the contention that the implicit measure is supported by explicit knowledge about the sequence structure (Perruchet & Amorium, 1992).

Subjective Awareness and Implicit Learning

In this experiment, subjects were asked if they noticed any regularity about the

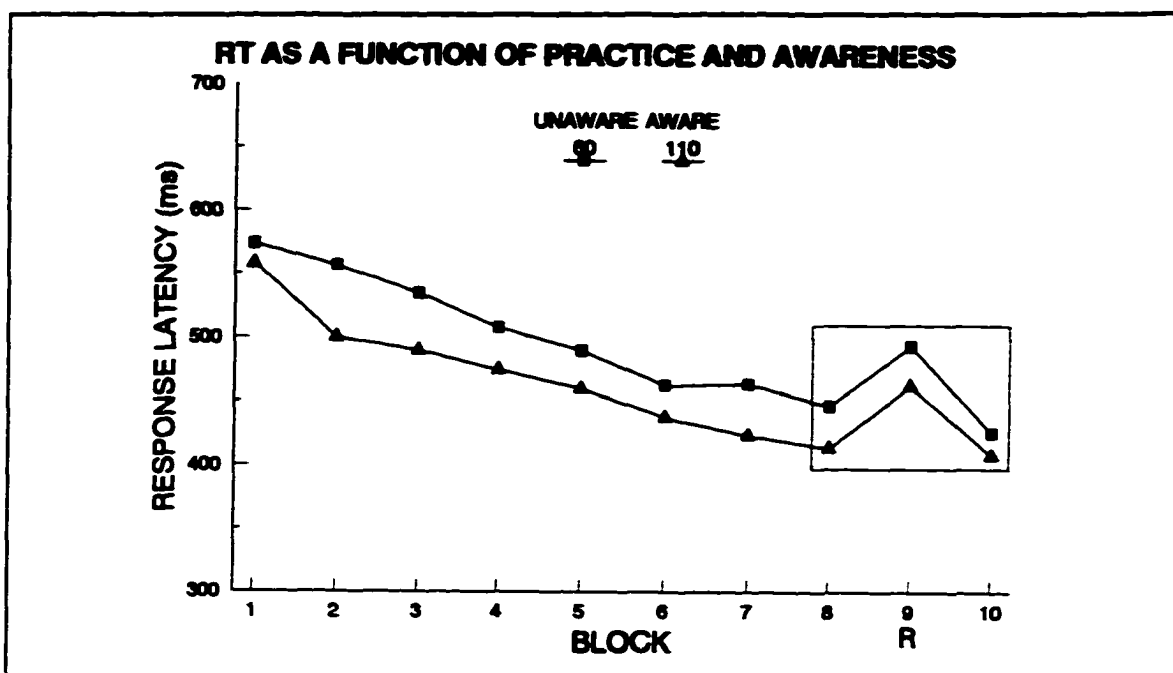


Figure 5.5 RTs are plotted across the trial blocks according to whether subjects were aware of the presence of a regular pattern.

event sequence after they completed the 10 trial blocks and before the start of the prediction task. Their answers were used to assess if they were aware of the presence of a pattern. This was intended to assess subjects' subjective experience of awareness. The resulting measure was then used as a grouping variable in a one-way ANOVA with implicit learning as the dependent variable. RTs grouped by this variable are shown in Fig. 5.5. ANOVA produced no evidence indicating that the "aware" subjects achieved more learning than the "unaware" subjects ($F_{(1, 168)} = .0418$, $MSe = 2169.7022$, $p = .8382$). This lends further support to the argument that the observed sequence learning is independent of conscious awareness. Of course, this result should be viewed with caution, for the conscious state could not be verified in an objective way when subjects claimed that they were aware/unaware of the presence of the sequence.

Exp. 2. Confidence Level and Their Relation with Sequence Learning

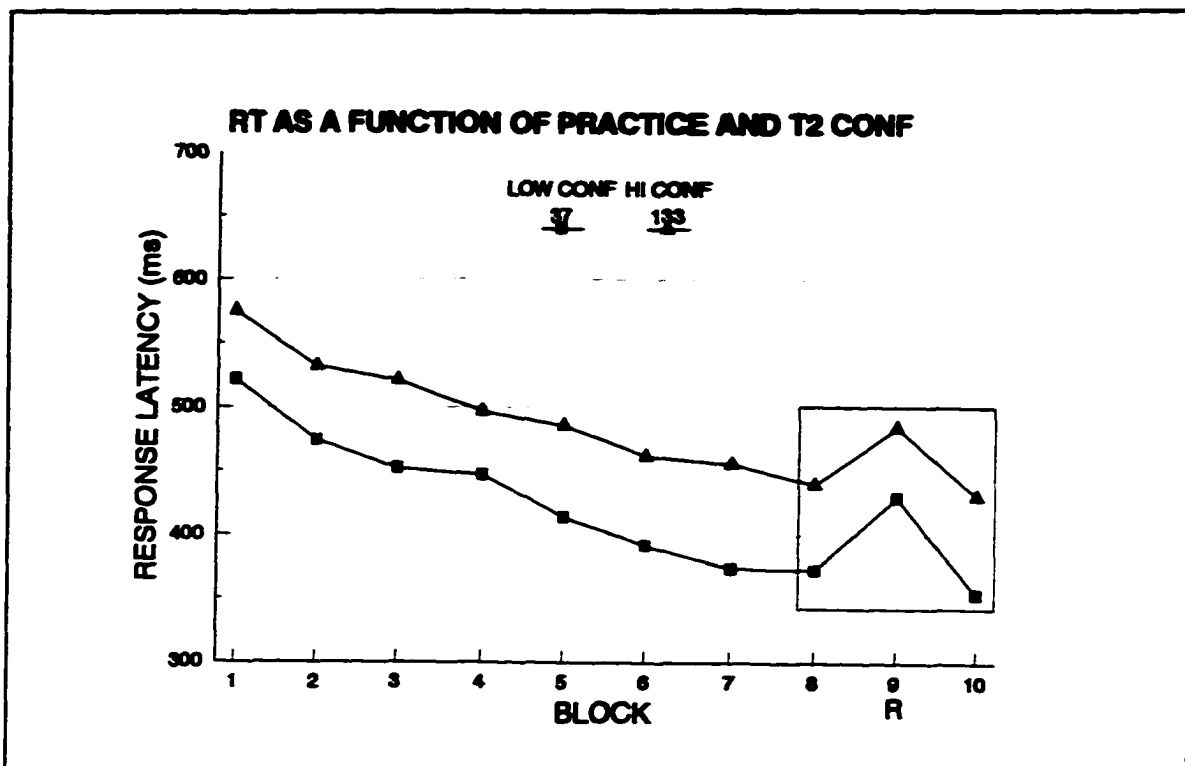


Figure 5.6 Relationship between sequence learning effect ($RT_7 - RT_6$) and confidence level (High vs. Low) for secondary task response.

At the end of each trial block of the SRT task, subjects were prompted to enter the number of times the target secondary signal occurred during that block. E_{R2} , which reflects a departure of subjects' input from the actual target count, was used as an dependent variable to check how well subjects performed the secondary task. They were also asked to indicate their confidence level on a scale between 1 and 7 (1 = least confident and 7 = most confident). For the present analysis, an average confidence level was calculated for each subjects. Confidence level between 1 and 3.5 was categorized as Low while greater than 3.5 up to 7 was categorized as High. Data are plotted in Fig. 5.6. Among many possible effects, the confidence rating task may have encouraged subjects to perform the secondary task with greater efforts even though no accuracy feedback was given in any block. Indeed, subjects' error rate

decreased markedly after the first block and leveled off after the second. Note that E_{22} was not affected by sequence change in the transfer blocks. Confidence level appears to be relevant only to secondary task performance. Although high confidence level was associated with slower RTs on the SRT task and vice versa ($F_{(1,160)} = 9.72$, $MSe = 127$, $p = .002$), it had no impact on sequence learning per se ($F_{(1,160)} < 1$).

5.4 Chapter Summary and Conclusion

This experiment varied both the amount of attention allocated to the secondary task and the stimulus onset of this distracting task. It was predicted that when subjects are paying due attention to the secondary task, and when the secondary task is presented in such a way that the processing of the secondary task overlaps with that of the primary task, sequence learning will be impaired to the greatest degree. Although there is some indication consistent with this hypothesis, it was not supported statistically. It is not clear why the expected results were not observed. This needs to be clarified in future work. In accord with the two sets of experiments reported before, sequence learning was observed, and it seems it was affected by how much attention was given to the distractor. But the support for the latter part of the statement was only marginal. Contrary to initial expectations, however, different tone-onsets failed to yield differential amounts of sequence learning even though this factor appeared to determine subjects' response rate. Nevertheless, crossing attentional allocation and tone-onsets did not produce sufficient evidence that availability of attention determines how much sequence learning can be achieved.

This experiment also attempted to address the implicit nature of sequence learning.

It was shown that subjective report of awareness, or the lack of it, is not predictive of the success or failure of sequence learning. Apart from that, the purported explicit measures and the implicit measure were poorly correlated, indicating that the two types of tasks do not share a common database. This brings us to the conclusion that implicit sequence learning can be achieved in the SRT paradigm with a distractor. What is amazing is how well the cognitive system can orchestrate the effective use of attentional resources by quickly shifting attention among several different tasks.

CHAPTER 6

GENERAL DISCUSSION AND CONCLUSIONS

6.1 A Summary of the Findings from This Study

Does implicit learning require attention? This is the question the present study attempted to address. The difficult part of the undertaking lies in the fact that the familiar concept of attention has never been satisfactorily defined. It is like fishing in a lake without knowing what exactly you are trying to catch and what you would expect to find. This being the case, you are never sure whether the tools you are using are appropriate for the task at hand. When Nissen and Bullemer (1987) first raised the question, they appeared to have conceived of attention as some sort of overhead processing with a limited capacity in terms of the number of tasks it can handle adequately at the same time. When they introduced the secondary tone-counting task, they thought this would engage that processing mechanism, thus forcing the SRT task to be performed when less attention is available. It sounds like a simple, straightforward idea. As the differences in performance measures between their repeating-pattern group and the random group failed to reach statistical significance, they concluded that sequence learning requires attention. However, the conclusion was challenged when later investigations employing different approaches (Cohen et al., 1990; Curran & Keele, 1993, Frensch et al. 1994) found, with some qualifications, that subjects were actually sensitive to the presence of a sequence. The

pendulum now swung in the direction suggesting that attention may not be necessary for sequence learning.

The positive findings just mentioned were based on sequences with at least some unique transitions in the training sequence. Some reservations were made by Cohen et al. (1990) concerning sequences consisted of ambiguous transitions. The present study showed that even sequences made up entirely of ambiguous transitions can still be learned. When interpreted, this finding will serve to remove the qualifications and lead to the conclusion that sequence learning may be achieved without attention. But, although bringing the pendulum to a full swing may be interesting in its own right, further analyses indicated that things are far more complicated than that. First, it was found that subjects who performed the secondary task relatively poorly showed much better learning than those who did well. If doing well on the secondary task implies paying more attention to it, this would suggest that sequence learning is affected by the availability of attention. The fact that giving more attention to the secondary task results in a compromise in sequence learning indicates that (1) the two tasks share common resources; and (2) these resources are limited in capacity. Second, it was also found that when subjects were given a practice on the secondary task prior to performing the dual-task SRT task, sequence learning was substantially reduced. It is likely that the practice showed the subjects how difficult it was to perform the secondary task adequately. This encouraged them to allocate more attention to the task, thus making attention less available to sequence encoding. However, analysis combining the data from Exps 2a and 2b failed to substantiate the difference due to the practice as was found when the data were analyzed separately. So the result

concerning the effect due to externally induced attention manipulation needs to be clarified by replication. Third, systematically manipulating the onset of the secondary task stimulus demonstrated that learning was impaired only when subjects were paying more attention to the secondary task and when the two competing tasks were presented in close succession. This suggests that the two cognitive processes can share the same resources successfully unless both tasks demand attention at the same time or in a very narrow time frame. Together, these findings invite the conclusion that attention remains a critical determinant in sequence learning.

6.2 Some Theoretical and Empirical Issues of Attention

In this study, as well as in many other studies on sequence learning, a secondary task was employed to manipulate the availability of attention to the processes involved in sequence encoding. The secondary task is called a distractor, which diverts attention away from other tasks. The presence or absence of a distractor determines the availability of attention empirically. But what is attention, anyway? To help us come to grips with the concept of attention, we must first understand that performing a task involves a series of high-level processes such as identifying a stimulus and selecting a response, which require various resources. Although psychologists do not yet have a clear understanding of these resources, one that has been extensively explored is the working memory. When a component of a task is given the necessary resources to be processed, it is made current in the system and becomes relevant here and now. But what determines which task receives what resources? There seems to be a central executive that assumes the role of such a function. This is what I would like to call attention.

I certainly do not intend to revive the old homunculus that sits atop overseeing the operations of the cognitive system. Rather, a neural network model is preferred to conceptualize the workings of a system with attentional functions. Attention, then, may be thought of as a set of "switches" driven by a summation of excitatory and inhibitory neural messages coming from a complex array of nuclei and circuits. These switches allow the prevailing message to be channelled to the appropriate resources. Some of the messages may come from activated memories of past experience, others from fresh sensory input. One interesting property of the attentional system is that it almost always gives priority to incoming sensory information. It even interrupts an on-going process and grant the resources to process a new stimulus. This of course makes perfect evolutionary sense. However, I would also think that external stimulation generates neural impulses of much greater intensity than those produced by spontaneous activities in the brain. As the attentional mechanism operates on a winner-take principle, this explains why external stimulus takes precedence over internal stimulus.

Coming back from a brief theorizing of attention, we have yet to explain why sequence learning is impaired under certain circumstances. In the context of the sequence learning experiment, a (primary) SRT task is performed concurrently with a secondary task that requires subjects to keep track of a target signal, two types of stimulus must be processed to the extent to which an appropriate response must be made to each. The first level of responses are to press a corresponding key and update or do-not-update the target count respectively. Processing at this level does not require a lot of a attention and can be achieved by quickly switching between the two

tasks. Initially, there may be some interference between the two tasks. But it disappears soon after automatization is attained. However, to actually update the target count requires subjects to retain the number last updated in working memory and perform an arithmetical operation, which requires substantially more resources for an extended period of time. No explicit instruction is given as to how the primary stimulus must be processed beyond producing a key-press. Encoding of structural properties of the event sequence is an evolutionarily equipped function which is carried out automatically without explicit instruction. Although one may not become conscious when the encoding process is going on, it does require resources as any operation.

There are several scenarios in which sequence encoding may be compromised or fail. A trivial one, for the purpose of the present study, is when the stimulus is random, or the statistical constraints of the events are weak (Nissen & Bullemer, 1987; Stader, 1992). A second one is when attention assigns priority in the use of the resources to other processes such as updating the secondary target count. This explains why subjects who performed a single task demonstrate better learning than those who performed a dual-task (Cohen et al. 1990; Nissen & Bullemer, 1987). A third one is when the secondary task is somehow emphasized so that attention allows more resources and for a longer period of time for the task. As was shown in Experiment 2b, subjects' error rate on the secondary task was significantly reduced when a practice session was introduced. But sequence learning was also reduced. The fourth one implies an interface of a conscious monitoring system with the attentional mechanism. When the secondary stimulus was presented immediately before the

primary stimulus, attention first gives resources to the first incoming request, which is from the secondary stimulus, then quickly to the second incoming request, which is from the primary stimulus. After that, consciousness takes control over attention and mandates that resources be yielded to updating the secondary target count. As a result, sequence learning is severely compromised as I have shown in Experiment 3. This brings us to a discussion on the relationship between attention and awareness.

6.3 The Relationship between Attention and Awareness

Primitive organisms interact with the external world in a strictly stimulus-response fashion. As such, they survive at the mercy of the environment. As more complex species were evolved, some acquired a monitoring function we call conscious awareness. This may or may not include what philosophers deem as the most essential component of consciousness, namely, the sense of self. An animal responding inappropriately gets gobbled up by its predator. It doesn't have another chance to respond differently after learning the consequence through conditioning. However, it will benefit from observing how other members of the species respond and which responses were correlated with ill fate. The inexperienced octopus learns to suck off the cork from a glass jar in order to retrieve some food by observing another octopus doing just that. This does not require that the organism understands that the incident involves an animal other than "me". What is necessary is that it encodes the situation and compares it with a copy of another similar situation and decides which is preferable. Taking the preferred course of action presupposes that the animal be aware that there exists more than one alternative and some alternatives correlate with better consequences.

In the absence of consciousness, the attentional system goes about its business routinely, determining which process receives resources according to the order of requests and intensity of the message. When the conscious monitoring device reports that a certain process takes precedence over others, attention responds by surrendering the processing resources to the preferred task. It is possible that the consciousness "nucleus" sends feedback to channels bearing the signals of the preferred task via excitatory connections. This amplifies the strengths of the original message, which overrides the attentional channel-guards, making resources available to the consciousness-favored task and sometimes purging out an on-going process prematurely. However, it must be noted that consciousness does not always interrupt the work of attention . It does so only when it "deems" necessary. This is why the Poor Counters and Estimators in the SSPSOA = 50 ms groups in Experiment 3 showed comparable levels of learning as their counterparts in the other two (100 and 150 ms) groups despite the fact that the primary task stimulus was presented immediately before the secondary stimulus. If subjects decide to belittle the task, or instructions implies it is not so important, consciousness does nothing to intervene in the business of the attentional mechanism.

References

- Cleeremans, A. (1993). *Attention and awareness in sequence learning*. In the Proceedings of the XVth Annual Conference of the Cognitive Science Society. Hillsdale, NJ: Erlbaum.
- Cleeremans, A., & McClelland, J.L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General*, 120, 235-253.
- Cohen, A., Ivry, R.I., & Keele, S.W. (1990). Attention and structure in sequence learning. *Journal of Experimental Psychology: Learning, Memory, and cognition*, 16, 17-30.
- Cohen, A., & Curran, T. (1993). On tasks, knowledge, correlations, and dissociations: Comments on Perruchet and Amorim (1992). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(6), 1431-1437.
- Curran T. & Keele, S.W. (1993). Attention and Nonattentional forms sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(1). 189-202.
- Eysenck, M.W. (1982). *Attention and Arousal: Cognition and Performance*. Berlin; New York: Springer-Verlag.
- Eysenck, M.W. (1995). *Cognitive Psychology: A Student's Handbook*, 3rd ed. Hillsdale, NJ: Erlbaum.
- Fagot, C., & Pashler, H. (1992). Making two responses to a single objects: Implications for the central attentional bottleneck. *Journal of Experimental Psychology: Human Perception and Performance*, 18(4), 1058-79.
- Frensch, P.A., Buchner, A., & Lin, J. (1994). Implicit learning of unique and

- ambiguous transitions in the presence and absence of a distractor task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 567-584.
- Frensch, P.A., Lin, J., & Buchner, A. (in press). Implicit sequence learning: Evidence against an attentional learning mechanism. *Psychological Research*.
- Frensch, P.A. & Miner, C.S. (1994). Effects of presentation rate and individual differences in short-term memory capacity on an indirect measure of serial learning. *Memory & Cognition*, 22(1), 95-110.
- Goschke, T. (1992). The role of attention in implicit learning of structured event sequences. Paper presented at the Annual Conference for Experimental Psychologists, Osnabrück, Germany.
- Hasher, L. & Zacks, R.T. (1984). Automatic processing of fundamental information. *American Psychologist*, 39, 1372-1388.
- Heuer, H., & Schmidtke, V. (1996) Secondary-task effects on sequence learning. *Psychological Research*, 59, 176-186.
- Howard, D.V., & Howard, J.H. Jr. (1989). Age differences in learning serial patterns: Direct and indirect measures. *Psychology and Aging*, 4, 357-364.
- Howard, J.H.Jr., Mutter, S.A., & Howard, D.V. (1992). Serial Pattern Learning by Event Observation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(5), 1029-1039.
- Hsiao, A.T. (1995). Modality effect in sequence learning under attentional manipulation. Paper presented at 66th Annual Meeting of the Eastern Psychological Association. Boston, Massachusetts.
- Hsiao, A.T. (1997). Implicit sequence learning depends on available residual

- attention. Paper presented at the 68th Annual Meeting of the Eastern Psychological Association. Washington, D.C.
- Hsiao, A.T., & Reber, R.S. (1996). Short (50-ms) distractor-target asynchronies disrupt SRT learning. Paper presented at the 37th Psychonomic Society Annual Meeting, Chicago, Illinois.
- Jiménez, L., & Cleeremans, A. (1994). Direct and indirect measures of implicit learning. *Proceedings of the 16th Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum, 445-550.
- Jordan, M.I. (1990). Learning to articulate: Sequential networks and distal constraints. In Jeannerod, M. (ed.), *Attention and Performance XIII*. Hillsdale, NJ: Erlbaum.
- Kahneman, D. (1973). *Attention and Effort*. London: Prentice-Hall.
- Keele, S.W. (1973). *Attention and Human Performance*. Pacific Palisades, CA: Goodyear Publishing.
- Keele, S.W., & Jennings, P.J. (1992). Attention in the representation of sequence: Experiment and theory. *Human Movement Studies*, 11, 125-138.
- Knopman, D. (1991). Unaware learning versus preserved learning in pharmacologic amnesia: Similarities and differences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(5), 1017-1029.
- Lee, Y. (1995). Effects of learning contexts on implicit and explicit learning. *Memory & Cognition*, 23(6), 723-734.
- Lewicki, P., Czyzewska, M., & Hoffman, H. (1987). Unconscious acquisition of complex procedural knowledge. *Journal of Experimental Psychology: Learning,*

- Memory, and Cognition*, 13, 523-530.
- Merikle, P.M. (1992). Perception without awareness. *American Psychologist*, 47(6), 792-795.
- Merikle, P.M. & Reingold, E.M. (1991). Comparing direct (explicit) and indirect (implicit) measures to study nonconscious memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 224-233.
- Millward, R.,B & Reber, A.S. (1972). Probability learning: Contingent-event sequences withlags. *American Journal of Psychology* , 85, 81-98.
- Myers, J.L. & Well, A.D. (1991). *Research Design & Statistical Analysis*. New York, NY: HarperCollins
- Nissen, M.J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance systems. *Cognitive Psychology*, 19, 1-32.
- Nissen, M.J., Knopman, D.S., & Schacter, D.L. (1987). Neurochemical dissociation of memory systems. *Neurology*, 37, 789-794.
- Pashler, H. (1989). Dissociations and dependencies between speed and accuracy: Evidence for a two-component theory of divided attention in simple tasks. *Cognitive Psychology*, 21, 465-514.
- Pashler, H. (1993). Doing two things at the same time. *American Scientist*, 81, 48-55.
- Pashler, H. (1994). Graded capacity-sharing in dual-task interference? *Journal of Experimental Psychology: Human Perception and Performance*, 20(2), 330-342.
- Pashler, H., & Carrier, M. (1996). Structures, processes and the flow of information. In E. Bjork & R. Bjork (Eds.) *Handbook of Perception & Cognition. Memory* (2nd ed.). San Diego, CA: Academic Press.

- Perruchet, P., & Amorim, M. (1992). Conscious knowledge and changes in performance in sequence learning: Evidence against dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 785-800.
- Reber, A.S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, *118*(3), 219-235.
- Reber, A.S. (1993). *Implicit Learning and Tacit Knowledge: An Essay on the Cognitive Unconscious*. New York, Oxford Univ. Press.
- Reber, A.S. (1997). Implicit ruminations. *Psychonomic Bulletin & Review*, *4*(1), 45-55.
- Reber, A.S., Allen, R., & Regan, S. (1985). Syntactical learning and judgment, still unconscious and still abstract: Comment on Dulany, Carlson, and Dewey. *Journal of Experimental Psychology: General*, *114*, 17-24.
- Reber, A.S., Allen, R. & Reber, P.J. (in press). Implicit and explicit learning. To appear in R. Sternberg (Ed.) *The concept of cognition*. Cambridge, MA: MIT Press.
- Reed, J., & Johnson, P. (1994). Assessing implicit learning with indirect tests: Determining what is learned about sequence structure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(3), 585-94.
- Schacter, D.L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 501-518.
- Schmidtke, V., & Heuer, H. (in press). Task integration as a factor in secondary-task effect on sequence learning. *Psychological Research*.
- Shanks, D.R. & St. John, M.F. (1994). Characteristics of dissociable learning

- systems. *Brain & Behavioral Science*, 17, 367-447.
- Servan-Schreiber, E., & Anderson, J.R. (1990). Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(4), 592-608.
- Stadler, M.A. (1992). Statistical structure and implicit serial learning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18(2), 318-327.
- Stadler, M.A. (1993). Implicit serial learning: Questions inspired by Hebb (1961). *Memory & Cognition*, 21(6), 819-827.
- Stadler, M.A. (1995). Role of attention in implicit learning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 21(3), 674-685.
- Stadler, M.A., & Neely, C.B. (in press). Effects of sequence length and structure on implicit serial learning. *Psychological Research*.
- Tipper, S.P., & Cranston, M. (1985). Selective attention and priming: Inhibitory and facilitory effects of ignored primes. *Quarterly Journal of Experimental Psychology*, 37A, 591-611.
- Tipper, S.P., Weaver, B., & Watson, F.L. ((1996). Inhibition of return to successively cued locations: Commentary on Pratt and Abrame. *Journal of Experimental Psychology: Human Perception & Performance*, 22(5), 1289-1293.
- Treisman, A.M. (1969). Strategies and models of selective attention. *Psychological Review*, 76, 282-299.
- Treisman, A.M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97-136.
- Treisman, A.M., Sykes, M., & Gelade, G. (1977). Selective attention and stimulus

integration. In S. Dornič (Ed.) *Attention and Performance VI* (pp. 333-362), Hillsdale, NJ: Erlbaum.

Waugh, N.C., Fozard, J.L., Talland, G.A., & Erwin D.e. ((1973). Age, stimulus repetition, and reaction time. *Journal of Gerontology*, 28, 466-470.

Willingham, D.B. (1993). Implicit sequence learning is disrupted by a within-task attention manipulation. Manuscript submitted for publication.

Willingham, D.B., Greenberg, A.R., & Thomas, R.C. (1997). Response-to-stimulus interval does not affect implicit motor sequence learning, but does affect performance. *Memory & Cognition*, 25(4), 534-542.

Willingham, D.B., Nissen, M.J., & Bullemer, P. (1989). On the development of procedural knowledge. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15(6), 1047-1060.