Order Number 9304700

On the representation of implicitly acquired knowledge: Transfer across stimulus forms and modalities

Manza, Louis, Ph.D.

City University of New York, 1992

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

A

ON THE REPRESENTATION OF IMPLICITLY ACQUIRED KNOWLEDGE: TRANSFER ACROSS STIMULUS FORMS AND MODALITIES

by

LOUIS MANZA

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.

1992

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.

July 2, 1992

Date

Chair of Examining Committee

July 1, 1992

Date

Executive Officer

Professor Matthew Erdelyi

Professor R. Glen Hass

Professor Neil Macmillan

Professor Don Scarborough

Supervisory Committee

The City University of New York

Abstract

ON THE REPRESENTATION OF IMPLICITLY ACQUIRED KNOWLEDGE:
TRANSFER ACROSS STIMULUS FORMS AND MODALITIES

by

LOUIS MANZA

Advisor: Professor Arthur S. Reber

Implicit learning (IL) is a process by which knowledge is acquired largely independent of consciousness of either the process of acquisition or the information ultimately acquired (Reber, 1989). Within the study of IL, a controversy has recently emerged over the issues concerning whether tacit knowledge is represented in either an exemplar, a fragmentary, or an abstractive manner. To determine which of these positions most accurately reflects the form in which tacit knowledge is represented within the mind, four experiments were conducted.

A transfer design was employed to address the implicit knowledge representation issue. This technique basically involved initially presenting to subjects materials that they were required to learn, followed by a test of their acquired knowledge. The power of the transfer design originated from the manner in which the form(s) of the information presented during learning and testing were manipulated independent of each other. What permitted this procedure

to investigate <u>nonconscious</u> aspects of learning is that the transfer task was implemented within a procedure, common to IL, referred to as an artificial grammar (AG) learning task. This task involved two phases, the first of which centered on the memorization of presented stimuli which were generated from a finite-state artificial grammar. During the second phase, subjects decided if presented stimuli adhered to the same rules that the memorized items followed. The aspect of this task that allowed insight into knowledge representation was that the training (Control) and testing (Transfer) items differed in their surface forms but retained basic underlying, structural commonalities.

The first two studies investigated transfer across different orthographic forms, by utilizing a between- (Experiment 1) and within- (Experiment 2) subjects design. The latter two studies, on the other hand, focused on modality effects and implicit learning, as subjects in these investigations had to attempt to transfer sensory information across modality-laden orthographic (Experiment 3), as well as spatial and temporal (Experiment 4) forms of stimuli. The hypotheses for the four studies centered on determining if tacit knowledge is stored in (1) solely an abstract, exemplar-based, or fragmentary form, or if nonconscious information is (2) more likely stored as an eclectic mix of these three data forms.

Although various dependent measures were taken, two have a direct bearing on the representation issue: d' scores and reaction times (RT's). Measuring subjects' sensitivities to Transfer and

Control stimuli indicates both the degree to which transfer does or does not occur and subjects' ability to classify stimuli of different surface forms. Response latencies are relevant to the representation issue due to the fact that they provide insight into the possible mechanisms of transfer. Brooks and Vokey (1991) have suggested that nonconscious transfer occurs by subjects making analogies between Transfer and Control symbols, a process which relies heavily on exemplar-based knowledge. Reber (1989) and Mathews (1991), however, hold that the transfer of implicit data occurs via the utilization of abstract knowledge. It is argued that exemplar-based transfer should yield longer decision latencies than abstract transfer, which highlights the notion that RT's can provide a wealth of information into how tacit knowledge is represented.

Experiment 1 revealed that subjects could transfer complex knowledge across different orthographic forms, although subjects' classification ability was better for Control items, in comparison to Transfer items. Reaction times for Transfer items were slower than Control RT's during an initial testing block, but this difference was virtually eliminated over time as Transfer and Control RT's merged closer together. This finding suggests that exemplar-based elements may play a vital role in the early stages of the development of a tacit system, although abstract elements assume a dominant role at a later time. Exposing subjects to an additional number of training and classification items for Experiment 2 (as compared to the number of items in Experiment 1) did not change the higher sensitivity for

Control items found in Experiment 1, although Transfer and Control decision latencies were equal in Experiment 2. Adding to these findings were the results from Experiment 3, which found better classification performance for visual stimuli, although positive nonconscious transfer was found to exist across both visual and auditory forms of items. This finding was replicated in Experiment 4, which also found equal RT's to Transfer and Control items, a finding which supports the earlier notion that while distributive data may be predominant during the initial stages of a representation's formation, abstract knowledge may assume a more vital role at later stages in a tacit memorial system.

Overall, these results are discussed in relation to how they correspond to existing frameworks of tacit knowledge, in addition to how these results may be applied to a new theoretical model of the implicit learning process.

Acknowledgements

The completion of a dissertation, while credited to an individual, is by no means an individual effort. I would like to use this space to thank all of those who have helped me over the last four years.

First, and most importantly, I want to thank my wife, Lisa, for all the love and support she has given me, as well as for all of the sacrificies she made to allow me to obtain my doctorate. I can never begin to repay her for all she has done; without her I would not have finished this dissertation in the short time that I did.

On an academic level, I would first like to thank my advisor, Arthur Reber, for all the help he has provided. His endless assistance, in addition to his willingness to allow me to think for myself, is greatly appreciated. I also wish to thank those faculty members at Brooklyn College who served on my dissertation committee: Professors Erdelyi, Hass, Macmillan, and Scarborough. Their insights allowed me to better understand my own work, as well as provide some impetus for future studies.

I am also grateful to all of my fellow lab members, Toni Genovese, Andrew Hsiao, Dr. Michael Kushner, Max Locke, Faye Walkenfeld, & Bill Winter, for participating in various discussions about some of the issues raised in this dissertation.

I would like to thank my parents, Diane & Jerry, and my Grandma Sue, for all they've done.

Finally, I would like to thank Johnathan and Nicholas, my two little guys, for making all the hard work worthwhile.

Table of Contents

List of Tables

List of Figures

Introduction

Imagine that two college students enter a psychological laboratory to participate in a study on memory. After being seated, each student is given a story to read. Subjects in Group 1 read a passage as follows:

> "Suppose you are a doctor faced with a patient who has a malignant tumor in his stomach. It is impossible to operate on the patient, but unless the tumor is destroyed the patient will die. There is a kind of ray that can be used to destroy the tumor. If the rays reach the tumor all at once at a sufficiently high intensity, the tumor will be destroyed. Unfortunately, at this intensity the healthy tissue that the rays pass through on the way to the tumor will also be destroyed. At lower intensities the rays are harmless to healthy tissue, but they will not affect the tumor either. What type of procedure might be used to destroy the tumor with the rays, and at the same time avoid destroying the healthy tissue" (Gick & Holyoak, 1980, pp. 307-308).

Once this passage has been read, subjects in Group 1 are asked to determine how the problem could be solved, yet they have much difficulty in doing this. Subjects in a second group, however, are given the following problem, with its solution, before being presented with the radiation problem:

"A small country fell under the iron rule of a dictator. The dictator ruled the country from a strong fortress. The fortress was situated in the middle of the country, surrounded by farms and villages. Many roads radiated outward from the fortress like spokes on a wheel. A great general arose who raised a large army at the border and vowed to capture the fortress and free the country of the dictator. The general knew that if his entire army could attack the fortress at once it could be captured. His troops were poised at the head of one of the roads leading to the fortress, ready to attack. However, a spy brought the general a disturbing report. The ruthless dictator had planted mines on each of the roads. The mines were set so that small bodies of men could pass over them safely, since the dictator needed to be able to move troops and workers to and from the fortress. However, any large force would detonate the mines. Not only would this blow up the road and render it impassable, but the dictator would then destroy many villages in retaliation. A full-scale direct attack on the fortress therefore appeared impossible.

The general, however, was undaunted. He divided his army up into small groups and dispatched each group to the head of a different road. When all was ready he gave

> the signal, and each group charged down a different road. All of the small groups passed safely over the mines, and the army then attacked the fortress in full strength. In this way, the general was able to capture the fortress and overthrow the dictator" (Gick & Holyoak, 1980, p. 351).

These two problems may be superficially different from one another, but structurally, they are identical. That is, generally speaking, the initial, intermediate and goal steps, as well as the solutions, of the two problem are identical. The typical result from this procedure is that subjects in Group 2 solve the radiation problem faster and with more accuracy than those in Group 1 (Gick & Holyoak, 1980, 1983). This result occurs because Group 2 subjects are: (1) able to notice the analogies between the original and test problem, and (2) access the general solution to the original problem, and hence are able to solve the new problem. Those who fail to solve the radiation problem, such as Group 1 subjects, in this setting, do so because they lack an appropriate source problem which would help them. This source problem, such as the one presented to subjects in Group 2, enables subjects to encode the relevant structural elements of the initial problem, and as a result, when they attempt to retrieve the solution, they can solve the problem more successfully than those who do not possess such knowledge (Gick & Holyoak, 1980, 1983). Studies such as this have been carried out for many years, and represent classic transfer studies. The ways in which subjects solve the various problems presented to them allows researchers valuable insight into how knowledge is represented within the mind.

Transfer Studies and Knowledge Representation

Overview

Transfer studies investigate how knowledge acquired on one task effects performance on a superficially different, but structurally similar, subsequent task (Ellis, 1965). By looking at subjects' performance on a transfer task, one can begin to speculate on the deep nature of knowledge representation, because for transfer to occur, one must have a comprehensive understanding of a stimulus environment (Singley & Anderson, 1989).

Although there are a variety of experimental designs that can be used to investigate transfer (see Ellis, 1965, for a review of these different designs), one design in particular sheds a great deal of light on the content of one's knowledge base concerning the domain under investigation. This transfer design is as follows. A control group is initially trained on a certain task (Task A), and is then tested on another task (Task A*) which uses the same stimuli as in Task A but in a different application. In addition, an experimental group also receives training on Task A, but during testing, they receive Task B, where the stimuli are physically different but structurally identical to those items in Task A*. Transfer is typically easily obtained in the control situation (Task A --> Task A*), but for positive transfer to occur in the experimental group (Task A --> Task B), one must engage in what has come to be known as transfer appropriate processing, or TAP (Graf & Ryan, 1990).

Those arguing in favor of TAP state that transfer is facilitated when the same set of cognitive operations are engaged during the initial and transfer tasks in these experiments (Graf & Ryan, 1990). From this perspective, assuming that a subject is not made aware of any subsequent testing task, if transfer occurs, the information on the learning task would most likely be represented in a domain-free form (Novick, 1990). It is therefore a widely held belief that in order for one task to facilitate performance on another task, one must develop a representation that contains commonalities between the two domains (Thorndike, 1925; Katona, 1940; Ellis, 1965; Singley & Anderson, 1989). Therefore, if positive transfer ensues, one can assume that a general, deep, representation of the domain in question has been developed. However, if performance on a transfer task is hampered by the original task, then one can assume that an individual has represented the information from each task in a domain-specific fashion. With these notions in mind, one can see how transfer studies and models may be able to provide a deeper insight into the nature of mental content.

Empirical Findings

Early studies. There exists a data base covering over 75 years of research exploring the existence of a relationship between transfer tasks and knowledge representation. The earliest transfer studies were concerned with the solving of arithmetic problems given previous exposure to similar problems. In a series of experiments, Thorndike and his associates (Thorndike & Woodworth, 1901; Thorndike, 1922, 1925) found that when concrete, stimulus-response, elements remain constant across tasks, transfer would result. While this approach, known as the Theory of Identical Elements, adequately accommodated Thorndike's data, it is not a comprehensive model. The major flaw of this theory is that it claims that any manipulation of surface features of stimuli would cause the transfer process to break down. This claim was based on the empirical finding that previous exposure to algebra problems did not facilitate performance on structurally identical but superficially different word problems (Thorndike, 1922). In this experiment, for example, subjects were initially shown and asked to solve the problem "X=6 and Y=23, and A=X + Y. What is A?" and then later given "B1=12 and B2=7, and C=B1 + B2. What is C?". The proportion correct for the initial problem was .94, while performance on the second problem was only .72. According to Thorndike, transfer did not occur in this situation because the specific stimulus elements in each problem were not identical. Later work by others (Katona, 1940; Gick & Holyoak, 1983, for example), however, showed that in order to increase the probability of the occurrence of transfer, multiple training problems should be given to subjects, something Thorndike failed to consider.

In this light, Katona (1940) conducted a series of experiments where he presented subjects with sticks arranged in certain geometrical configurations, and instructed them to arrange them in a different, specified, configuration. In one study, one group (CON) of subjects was trained to arrange the sticks in one configuration, but another group (MEM) was trained with multiple configurations, using different initial and final arrangements. When given a test problem (which was identical to the problem that CON subjects were trained with), the CON subjects solved it faster and more often than the MEM subjects. However when additional, novel problems were then presented to both groups, the MEM subjects were able to solve each problem more often and with shorter latencies than the CON group. This result supported the idea that conceptually related multiple-exemplar training leads to a more general representation of the stimulus environment. This hypothesis has since been tested and validated numerous times by modern researchers, with the general conclusion being that transfer is fostered by a representation that contains general, abstract components, rather than a multitude of specific facts about a specific domain of interest (Gick & Holyoak, 1980, 1983; Reed, 1987; Fong & Nisbett, 1991).

<u>Modern studies</u>. Understanding the representations that develop through the transfer process involves an understanding of the mechanisms of transfer. Two specific processes, encoding and retrieval, play a role in transfer, and how they operate can inform researchers as to how information is structured and processed. In

delineating encoding and retrieval processes as they relate to transfer, current approaches have determined that, at first, a problem or task must be encoded in such a way as to emphasize relevant structural components of a stimulus instead of superficial surface features (Gick & Holyoak, 1980, 1983; Gentner, 1983). As a result, transfer is said to be facilitated when one encodes information in a general, abstract, form (Fong & Nisbett, 1991). Encoding, however, is only the first step in the transfer process. In order for transfer to actually occur, one must apply previously acquired information from its original domain to another domain. This application occurs as a result of retrieval, in some form, of the encoded data. Transfer is truly facilitated, therefore, when source information is properly encoded into, <u>and</u> successfully retrieved from, memory, and then mapped onto the new task (Reed, 1987).

Transfer Models

As a result of the above, and other, empirical validations, it is now believed that transfer tasks are highly efficient in determining mental content (Singley & Anderson, 1989). Several accounts have been proposed to explain both the nature of the content of the mind and how transfer might occur within a knowledge representation (KR) system. In general, the occurrence of transfer points to hypotheses that knowledge is stored in an abstract, general form (Gick & Holyoak, 1983), similar to the probabilistic view, described shortly. How a KR structure might handle the transfer of this general knowledge has been proposed by several models. While there are slight differences between these approaches, they all share the emphasis on the mapping of general, structural features from one domain to another (Gick & Holyoak, 1983; Gentner, 1983). Holland, Holyoak, Nisbett, & Thagard (1986) argue that knowledge is represented in a production-like system (consisting of condition-action statements of the form 'If X Then Y'), and for transfer to occur, a new problem must be transformed in a way that makes it appear analogous to the original problem. Once this transformation occurs, activation will spread from the source analog to the target analog. When the target analog is activated, any mapping rules that exist within one's KR system are in turn activated, with the end result being the occurrence of analogical transfer. This idea of mapping is also the focus of Gentner's (1983) structure mapping model of transfer. However, in this case the source problem is not transformed; rather, surface attributes are simply removed at the time of retrieval, preserving the crucial structural relations that exist between two knowledge domains. These relations are then mapped from one domain to another, resulting in a hierarchical network of structural information. For example, one might study all there is to know about atoms, and then study the solar system. In understanding the relationship between these 2 domains, one might (1) retrieve the vital information of electrons rotating around a nucleus, (2) discard irrelevant specific information about electrons and nuclei, and (3) retain the basic notion of outside elements rotating about a central core. This final step has a direct correlate in studying the solar system, and since it is devoid of specific data, it can be applied to studying the relationship between the planets. However, while the idea of structure mapping is a plausible explanation concerning how knowledge is represented and transferred between domains, another account takes the position that concepts are stored in the form of a multi-dimensional schema (Gick & Holyoak, 1983) formed during encoding. Such a system primarily utilizes structural information, but some surface components also enter into the system. The nature of schemas, as well as production systems, will be discussed in more detail below.

Evaluation. There is no doubt that transfer studies and models provide insights into the nature of knowledge representations. However, traditionally speaking, research on knowledge representation has not focused on transfer studies. Investigations into the content and structure of the mind have used other methods to investigate internal mental processes, but the exclusion of transfer findings from these models makes them incomplete in explaining how knowledge is represented. The discussion will now turn to these perspectives, with the purpose of delineating how these accounts might incorporate transfer results, and how such a merger would effect the nature of these models.

Knowledge Representation: An Overview

Theories concerning the manner(s) in which individuals represent information were first proposed by such 'classical' philosophers as Aristotle and Berkeley (Pikas, 1966). While modern philosophers still wrestle with the questions surrounding knowledge representation, the past century has seen numerous psychologists, from Wilhelm Wundt and his ruminations on abstraction (Pikas, 1966), to modern artificial intelligence theorists (Anderson, 1983), postulate various frameworks and mechanisms to account for knowledge representation. Such models have primarily focused on the representation of explicit, consciously held knowledge. Recently, however, a debate has arisen within the field of implicit cognition concerning the form in which knowledge acquired largely outside of the realm of conscious awareness is represented. Regardless of the implicit or explicit status of information, in determining the nature of knowledge representations, two issues must be discussed: (1) what is the form of mental representations, i.e., what type of knowledge is stored (specific, general, or both types of information), and (2) what is the structure of the mind, i.e., what type of mental architecture 'houses' these cognitive contents. Before discussing the various approaches to nonconscious knowledge representation, a review of the general literature on KR is in order, since implicit learning accounts of knowledge representation are based, in part, on these earlier models. Once this initial review is completed, the discussion will then focus on the issues surrounding the representation of tacit knowledge, in an ultimate attempt to develop an approach which can adequately handle the representation of implicitly acquired information.

Over the past century, numerous KR theories have been proposed. For the present discussion, a distinction will be made between the 'classical' theories of KR and the more 'modern' approaches. Specifically, the first part of this discussion will briefly review the former models, while the remaining portion of this initial discussion will focus on the latter theories. Across both domains, numerous theories have been proposed. The approach that will be taken towards this discussion of knowledge representation will therefore be one of general review of the area in question, followed by a brief evaluation of the approach.

Classical Approaches to Knowledge Representation

Two general models (concerned with lexical access and semantic memory) of knowledge representation, surfacing almost two decades ago, are no longer in vogue when it comes to theorizing about the structure of the mind. Current KR models, however, have evolved somewhat from these theories, so a brief discussion of these 'classical' views is appropriate at this juncture.

General Models. Lexical access theories have investigated the process whereby the understanding of individual words is accomplished either by isolating words or enclosing them within some type of sentential context. Models in this domain are formulated by interpreting results from procedures which may ask subjects question such as "Is HUMANE a word?" or "Determine if the word in capitals is a word: The boy was HUMANE to the dog." A general assumption of lexical access models has been that the speed with which a word is identified provides insight into how lexical data are organized mentally. In general, these approaches have proposed the manner in which the search for a word occurs. Overall, there seems to be a split decision concerning the nature of 'word searches', in that lexical searches are either: (1) serial in nature, starting with the most frequently occurring lexical forms (Becker, 1976; Forster, 1981), or (2) of the form where various word detectors attempt to identify different features (physical, sensory, etc.) of a word, combining their discoveries to facilitate lexical access (Morton, 1969, Stanovich & West, 1981).

Semantic memory approaches to KR, on the other hand, have concerned themselves with the representation of meaning (Chang, 1986). Here, stimuli are presented as "Is a canary a bird?" or "Is a bird an animal?" As with lexical access models, the speed at which meanings can be determined is postulated to provide information regarding memorial organization. However, these models view representation as more than just a listing of words; semantic models can be divided into two classes of hierarchy approaches: network and stage models. Network models postulate that information is stored by way of concept nodes connected by relational links (Collins & Quillian, 1969; Glass & Holyoak, 1974/1975). The general manner in which such network models operate is by a spreading of activation (Collins & Loftus, 1975) occurring from the presented concept to all related concepts, with a counter device recording the number of features or concepts encountered. Once a certain criterion is exceeded, the correct meaning can be determined. Network organizations are contrasted by stage models. The general stage approach is to first compare either both characteristic <u>and</u> defining features of the subject and predicate of a statement in a two-stage approach (Smith, Shoben & Rips, 1974), or just characteristic features in one step (McCloskey & Glucksberg, 1979). If enough features common to the subject are shared by the predicate, a positive match can be determined. However, if a decision cannot be reached during this first stage, the processing continues in a second stage, where all attributes must be shared in order for a positive match to result.

<u>Evaluation</u>. While these 'classical' models provided a starting point for modern KR theories, they are insufficient for explaining the organization of the mind because, among other reasons, they are too domain specific--knowledge is most likely represented in a variety of ways, not solely by the manners in which the 'classical' models propose. As far as transfer is concerned, it seems that for transfer to occur, some general knowledge is required. Since the lexical access models are based solely on specific words and how these words are stored in the mind, these models are highly content-specific. As a

result, transfer would be difficult to obtain in this setting. On the other hand, semantic approaches might be very well suited to handling transfer. Since these models contend that meaning is stored within a KR system, and transfer models dealing with knowledge representation also emphasize meaning over surface attributes, transfer tasks could add to the data base on semantic memory models. Overall, however, lexical access and semantic memory approaches have been eclipsed by more advanced theories of KR, and it is to these 'modern' approaches that this discussion now turns.

Modern Approaches to Knowledge Representation

Overview

It has been the opinion of numerous researchers (Pikas, 1966; Smith & Medin, 1981; Medin, 1989; Smith, 1989) that the mind is organized by relationships between conceptual information. Intuitively, this seems to be a most plausible characterization of knowledge representation, since it has been suggested that without conceptual structures, the mind would be a cluttered sea of ideas floating around with no sense of relation or purpose (Smith & Medin, 1981). With this in mind, virtually all current approaches to KR have concerned themselves with the organization of conceptual information. To understand the issues concerning conceptual processes, two issues, mentioned earlier, must be addressed, specifically a) the form that conceptual information assumes, and b) the structure that stores conceptual data. These two points will be considered in turn.

Several empirical approaches have been undertaken in an attempt to determine the content of conceptual representations. Although many different techniques have been employed in past empirical investigations, two methods are most common. One method initially involves subjects being presented with a general concept (e.g., dog), with their task being to list any and all ideas or instances associated with the presented concept. The common finding when utilizing this procedure is that more typical instances of a concept are retrieved before atypical ones, for example, a collie may be mentioned before a daschund (Rosch, 1978). The other common procedure reverses the former approach, in that subjects must decide if an instance belongs to a general concept (e.g., "Is a collie a dog?"). In this situation, typical instances (such as the previous example) are verified faster than atypical instances (e.g., "Is a daschund a dog?") (Smith, Shoben, & Rips, 1974). Various other procedures have been utilized in this vein, with the common finding being that typical instances of a concept seem to be accessed faster and more frequently than less typical instances (Smith, 1989). These typicality findings have been used as the data in the debate concerning the content of conceptual structures, a discussion which has resulted in the formation of three models concerning the content of conceptual structures.

Content Models

The Classical View. The first of these approaches is referred to as the Classical View (Smith & Medin, 1981). This perspective has focused on the abstraction of information from a stimulus environment (Bruner, Goodnow, & Austin, 1956; Pikas, 1966), with the abstraction process involving the disregarding of irrelevant surface attributes of a stimulus; attention is focused on the more general structural characteristics of stimuli. The result of this process is the formation of a single summary representation of a prototypical category member. For example, if one were to form an abstract prototype of a dog, one might include such vital features as 4 legs, less than 3 feet in height, and the ability to bark. Specific data such as hair color, temperament, odor, etc., are not necessarily vital to this prototype, since such secondary elements will most likely vary widely from one dog to another. Therefore, an abstract representation contains all features that an instance must possess to be considered a member of a particular concept, and if a subject is asked to determine if a presented instance belongs to a certain concept, the instance must possess all of the features of the prototype before being classified as a member of the concept. A main problem with this explanation, however, stems from the fact that the summary representation is formed primarily through contact with typical members of the concept, and all subsequent instances must possess all of the necessary and defining features of the prototype in order to be classified as an instance of the concept. This constraint causes problems when an atypical instance of a concept is encountered, because although it might contain some of the defining features of the prototype, some will most definitely be missing. Because of this emphasis on necessary and defining features, the Classical View has fallen out of favor as the best explanation for the content of the mind. While it may indeed be the case that general, abstract information is stored in a KR system, it is clearly not the only form of information that is represented within the mind. The next two approaches deal with these additional elements of representational systems.

The Probabilistic View. As discussed by Smith & Medin (1981), the Probabilistic View assumes the same notion of a summary representation being stored in one's memory about category members during initial acquisition, but the emphasis on necessary and defining features, central to the Classical View, is absent. Rather, the Probabilistic View assumes that the stored representation of a concept contains some aspect of central tendency that can account for a wide variety of instances, both highly and slightly similar to the prototype (e.g., DOG--4 legs, brown or black hair, barks, bad breath) (Smith, Shoben, & Rips, 1974; Collins & Loftus, 1975; McCloskey & Glucksberg, 1979; Medin, 1989). To handle the idea of central tendency, three 'sub-views' of the Probabilistic View have been developed in the literature (Smith & Medin, 1981). Although slightly different, the commonality that these approaches share is that they all postulate that the mind contains conceptual prototypes. The featural approach, for one, holds that the general conceptual information stored in one's memorial system is a set of discrete modal features which will be found in different degrees in different instances of a concept. In contrast, the dimensional approach holds that instances of concepts are represented along continuous dimensions. This differs from the featural approach in that dimensions are continuous, while features are discrete. Finally, the holistic approach views summary representations as being a type of template that accounts for an average of all features of a concept.

The Exemplar View. By looking at the first two views, one would think that the mind is simply a storage space for general information. However, the final content model, referred to as the Exemplar View (Brooks, 1978; Medin & Schaffer, 1978; Medin, 1989), eschews the idea of a summary representation in favor of a more detailed representational system which stores every example of a concept within a network-like structure. As instances of a concept are encountered, they are simply added to one's data base on a certain concept. In keeping with the above example, if one encounters a collie, all the characteristics that make it a collie are recorded. Then if one later encounters a beagle, all of its features would be recorded, separate from the elements of the collie. Of course, as common instances will tend to repeat themselves, these features will have the greatest chance of being associated with particular concepts. However, this view on the content of the mind is intuitively inefficient, as the repeated storage of identical conceptual instances would seemingly lead to a mind that was extremely cluttered.

Evaluation. These three approaches to the content of the mind offer several possible forms that knowledge may take within the mind. The question still remains, however, as to which of these perspectives is the most accurate. It has become clear in the literature (Medin, 1989) that the Classical View cannot account for the data supporting the other views, so that approach seems to be inappropriate. With the Probabilistic and Exemplar views remaining to account for the content of the mind, both positions make valid claims. The most plausible explanation seems to be that the information stored within a KR system is both general and specific. Information that is highly typical of a concept has a greater chance of being stored in a general form, while atypical instances of a concept are probably stored in a more specific form. The role of transfer within these three models is relatively clear. As before, with transfer relying on the representation of more general than specific data, the Classical and Probabilistic views are better suited to handling the transfer process. The Exemplar View, discussed later in more detail, seems to be intuitively less appealing for transfer, since this perspective is based upon the representation of specific events--data which the transfer process cannot handle as effectively as abstract information.

Although knowing the content of the mind is necessary for understanding how knowledge is represented, the possession of knowledge in and of itself means relatively nothing. Thought processes in general are concerned with how knowledge is applied, an aspect that the previous three views failed to enumerate. In order to gain a complete understanding of the content of the mind, and therefore how knowledge is represented, one must also know how knowledge is applied, because the application of knowledge is directly effected by both the content of the mind and the processing mechanisms involved in KR frameworks. The main approach used to investigate the application of knowledge has been the transfer of training technique, and it to this method of investigation that the discussion continues.

Frameworks of Knowledge Representation

Each of the following approaches attempting to understand the structure of the mind can be expressed in very detailed formats. The purpose of this review, however, is not centered on the intricacies of KR frameworks. Therefore, the approach to be taken with each framework will initially be one of general review. In addition, since the point has been made above that transfer studies have a direct tie-in to KR theories, each framework will be assessed in its ability to deal with the transfer of knowledge.

Instance-Based Structures

Overview. The simplest frameworks for storing and retrieving knowledge are perhaps those referred to as either instance-based (Brooks, 1978), context (Medin & Schaffer, 1978) or memory trace (Hintzman 1986) models of KR. Although the names of these models differ, they are all based upon the main assumption that what is stored in a KR system is very specific information, acquired through experience with the environment.

Models. The organization of such a model is basically one where instances are simply stored within a conceptual network that treats information as either an exact copy of an instance (Medin & Schaffer, 1978) or as a memory trace represented by the most primitive attributes of an experience (Hintzman, 1986). These models would seemingly take an exemplar approach to the content of the mind.

Within these instance based models, memory traces are stored until they are needed. Retrieval from a memory array occurs when a probe item is 'inserted' into the system and compared to instances that are similar to it. This process can occur by either a serial comparison of each instance to the probe (Medin & Schaffer, 1978) or by a simultaneous comparison between a summary of all related instances and the probe (Hintzman, 1986). However the process occurs, identification seems to be highly dependent on a similarity function.

Evaluation. Although memory-based approaches to KR may be simple to understand, their simplicity is their main drawback. While concept verification may indeed occur via some type of similarity heuristic, an internal structure based upon an instance-based notion would appear to be loosely organized and overly cluttered with specific information. Such a reliance on specific data would seemingly lend these models to be ineffective in transferring knowledge, since it seems to be the case that the transfer process depends on an abundance of general information to occur effectively (Singley & Anderson, 1989). However, Brooks and Vokey (1991) have recently hypothesized that instance-based KR structures can transfer knowledge by making similarity analogies between different domains. Their results, described in more detail later, do not entirely support this hypothesis, and to preview the current results, such a mechanism was found to provide only a partial explanation of the transfer process. Therefore, although these instance-based models do have many good points, as an overall architecture of the mind they lack an efficient mode of processing information.

Schema Theories

Overview. Another possible structure that may store conceptual information is what is known as a schema. The exact nature of what a schema consists of has been debated since the 1930's (Bartlett, 1932) but to date there has yet to be any firm conclusions on this topic (Alba & Hasher, 1983). While the notion of a schema is relatively old, it has recently come back as being a possible framework for knowledge representation (Alba & Hasher, 1983; Smith, 1989). In general, the term 'schema' is used to refer to a general knowledge base that an individual possesses about a specific domain of interest (Alba & Hasher, 1983), e.g., one's knowledge about automobiles. Empirical studies usually concern themselves with how individuals use and/or modify various knowledge schemas, according to five basic encoding processes--selection, abstraction, interpretation, integration, and reconstruction.

Models. Two thoroughly systematic schema theories have been discussed in the literature on this topic, and both of them assert that schemas consist of structural knowledge about various domains of interest to individuals. Frame theory (Minsky, 1975) asserts that knowledge representation is accomplished through the use of frames, the content of which are general expected information about events--no specific aspects, except for the ordering of events, are stored. Schank & Abelson's script theory (1977), on the other hand, postulates that scripts contain the same general information as frames, but they also contain specific information about the events they represent. For example, if I were to have a schema about driving, it would include basic information such as how to start, shift, steer, and stop the automobile. This would make up the frame. My driving script, however, while possessing the same information as in the frame, would also contain data about driving in Staten Island, New York (where I live). Such additional information would include watching out for reckless drivers, knowing that individuals will not signal before turning, etc. Regardless of the form of knowledge contained within a schema, there is a certain organization to them, as well as certain processes that oversee the information within them. An understanding of these functions is crucial to the development of schema models of KR.

Organization. Smith (1989) has offered an analysis of the structure of schemas, which are postulated to consist of five factors.

First, and most basic, is the notion that a schema consists of different general attributes of a concept (e.g., CAR--body size, transmission, fuel economy, engine size, etc.). Then, each attribute possesses several possible instances that it may assume (e.g., body size--compact, midsize, luxury). One of these instances is noted as a default, i.e., an instance that the attribute assumes to take on if no other information is given. This default is similar to the idea of a prototype suggested by the probabilistic view discussed above. These instances are also grouped together under one 'superset' that denotes the general concept that the individual schema is a member of (e.g., TRANSPORTATION VEHICLES: Car--Bus--Train--Bicycle--Airplane). This allows specific schemas to be linked to other schemas within the same concept. Like the semantic network models mentioned earlier, schemas also have relational links connecting attributes that are related to one another. Lastly, an attribute is marked according to its importance to an understanding of the entire schema. When attempting verification, described below, the more vital attributes would be checked first to judge concept membership. With this general framework in place, it is now possible to see how information is stored in a schema.

Processing mechanisms. The processes of selection, abstraction, interpretation, integration, and reconstruction are by no means limited to the understanding and/or explanation of schema theories; many other cognitive approaches utilize these mechanisms. However, these processes, as a group, are critical to the schema approach to KR (Alba & Hasher, 1983).

When one first encounters information to be processed, the message must be attended to. In order to access a relevant schema, certain aspects of the presented stimulus must be selected and applied to the schema to see if the new information can become part of the existing knowledge framework. Whether or not a presented message is selected to be encoded within an existing schema hinges on three factors: (1) whether or not a schema relevant to the input message exists, (2) whether or not a relevant schema can be activated, and (3) whether or not the new information is important or vital to the activated schema. If all of these criteria are met, the abstraction process begins. If not, the information is either discarded or applied to the formation of a new schema.

Once relevant new information is selected for encoding, any superficial surface information is removed from the message by the abstraction process. This process retains the meaning of the message while 'stripping away' the original format. In order to make any sense out of the abstracted form of the message, one must attempt to ascertain the meaning of the structure of the information. This is accomplished through interpretation, or inference, into the meaning, and there are two forms that have been proposed (Alba & Hasher, 1983). Pragmatic implication involves the conversion of the explicit message into its most likely "deep", or underlying, meaning. The other form of inference occurs during the comprehension of vague information, and is referred to as instantiation. During this process, general information is converted into more specific information, and that knowledge is stored in memory. Another form of instantiation occurs when certain aspects of a piece of information are missing. In this case, an individual will utilize the context surrounding the message to interpret the meaning of the entire message. Once a message has passed through the previous three processes, it becomes part of an existing schema (or is formed into a new schema) by way of the process known as integration. This occurs by combining whatever information has been selected, abstracted, and interpreted from a presented message and fitting it into a schema, where the new and/or existing framework may be modified in order to create a unified, holistic representation.

The final process involved in the schema process is known as reconstruction. This process takes place when one attempts to reproduce an episode stored in memory, and involves taking any episodes of a schema that are accessible in memory and combining them with general knowledge about the schema, in an attempt to re-create the schema from memory. Since this process involves some fabrication on the part of the individual, many errors between the reconstructed and actual schema may occur.

Evaluation. The schema view of knowledge representation is indeed a comprehensive and possible explanation of how representations are originally formed and subsequently structured. The inclusion of learning mechanisms is vital to the understanding of knowledge representation, and this inclusion makes schema theories plausible explanations of mental activity. As for the ability of a schema to handle the transfer of knowledge, such structures seem better suited for transfer due to the fact that schemas can hold general information. Indeed, the Gick & Holyoak (1983) model of knowledge representation is derived from transfer studies, with the model's main structure being a schema which contains general information about solving problems. However, although somewhat effective at explaining mental structure, there remain other ways to postulate the mental architecture, in addition to schemas.

Production System Models

Overview. Theories based on production systems have come into vogue over the course of the past 10 years or so, coinciding with the computer "explosion" that has occurred worldwide. In their most basic form, these systems (Anderson, 1983) are based on the fundamental idea that the mechanisms behind human cognition are vast sets of condition-action statements referred to as productions (e.g., IF little children are required to be immunized against polio and only doctors administer the immunizations, THEN a child must go to a doctor to be immunized.). When information is presented to a production system, the condition element searches for a match in short-term (or working) memory. If a match is found, the action is carried out; conversely, if the search fails to find a match, the action is terminated. Numerous production systems have been proposed to account for how this process operates, but two in particular exemplify the nature of the field--the ACT* approach (Anderson, 1983) and the SOAR model (Laird, Newell, & Rosenbloom, 1987). To gain a better understanding of how production systems propose to represent knowledge and thought processes, each of these models will be summarized below.

Models. The ACT* (Adaptive Control of Thought) model of cognition (Anderson, 1983) is made up of three types of memory systems: working, declarative, and production. Working memory consists of information that has either been recently presented in some sensory fashion or has been retrieved from the declarative store. Declarative memory is made up of facts about the world that have been encountered and encoded at some previous time, and production memory consists of production rules of the type described above. Within this framework, three types of knowledge representations operate. Temporal KR's code the orders of sets of information, spatial KR's are concerned with the spatial arrangement of information, and abstract KR's primarily encode the meaning of stimuli. In general, these three types of representations interact within the ACT* framework in the following steps. Sensory information from the environment is put into working memory. This information is then either stored into declarative memory (at which point the process ends), or is used as a "probe" to search declarative memory for a match in order to carry out a production. If a match is retrieved, this match is applied to production memory, where an additional match is searched for. If another match is found, the production is executed into working memory, and the process ends (or begins again if several productions are necessary to carry out an action).

Although any of the representational systems, production or otherwise, described so far can be characterized as 'problem-solvers', the SOAR architecture is specifically designed to solve problems of various natures (Laird, et al., 1987). The SOAR (State, Operator, And Result) approach views any situation as a problem that needs to be solved. The solution, or goal, is attained by applying various productions in a stepwise fashion through a series of subgoals, continuing until the desired goal is attained. At this point, the SOAR model is almost identical to the ACT* method of applying productions. The two approaches differ, however, in the fact that whereas ACT* has three types of memories and representations, SOAR has a single framework for long-term knowledge, that structure being a production system.

The type of production system that SOAR utilizes is one that fires all satisfied productions in a parallel fashion. When a problem has to be solved, the entire problem state is represented in working memory. Within working memory, there exists a 'context stack' which determines the hierarchy of problems and goals, objects (which are the states and goals themselves), and preferences which play a role in the encoding of procedural knowledge. Working memory is connected to a processing system, which contains production memory and a decision procedure. Production memory is simply a storage system of the actual productions necessary to solve any problem, and the decision procedure analyzes the information currently in working memory in order to guide the problem solving procedure. Connected to the production system is a chunking mechanism which adds new productions to production memory by constructing them out of existing productions. Finally, there is a working-memory manager which removes used or non-necessary information from working memory in order to streamline the problem-solving process.

Evaluation. These overviews of the ACT* and SOAR models are very simplified. The models, in their full formulations, are exceedingly complex, and a full explication of their functions is outside the scope of this discussion. The attempt here was to provide additional examples of how knowledge may be represented in the mind. To their credit, these production approaches take the most comprehensive and unified view of knowledge representation, encompassing virtually all of the previous approaches. The one drawback to these systems, however is that the knowledge contained within them is very domain-specific. While this is fine for dealing with specific problems, it presents a problem when attempting transfer. In order for transfer to occur, one must represent knowledge in a general form, devoid of much specific information. The only way that a production system could accomplish this is if there were some mechanism built into the system that could abstract vital information and discard irrelevant data. Such a model, to the best of my knowledge, does not yet exist, leading to the conclusion that while production systems are effective at modeling human thought, they are not the most effective structures. Such a complete framework does exist, however, and it appears to be perhaps the most comprehensive of all approaches to knowledge representation--connectionist networks.

The Connectionist Approach

Overview. While production systems model the architecture of the mind as an analogy to a computer, connectionist frameworks consider the representation of knowledge as the brain itself would function. This is, however, an assumption that has not yet borne out. As a result of this neural approach to knowledge representation, connectionist models are exceedingly complex, but they are also the most comprehensive KR frameworks developed to date.

Mechanisms. Connectionist models (McClelland & Rumelhart, 1985, for example), in their neural approach to KR, utilize a multitude of processing units, resembling an abstraction of a neuron, as the core of any KR system (Fodor & Pylyshyn, 1988; Rumelhart, 1989). These units are arranged in a layer-like fashion (Hanson & Burr, 1990), with thought represented as interactions between different units at the same and/or different levels. While the actual processing that produces mental activity is complex, any connectionist model can be understood by looking at several vital components (Rumelhart, 1989).

The basic processing component, as mentioned above, is referred to as a unit. Each unit within a connectionist framework can store bits of various information about an object, and interact with other related units to produce some type of meaningful output. A second assumption of a connectionist network is that there is a certain level of activation which serves as a pattern of activity over all units contributing to a representation at a given time. To achieve this type of activation, each unit has an output function which determines whether or not a connected unit will fire or inhibit the data feature that it contains. For example, if a compact car is a very common type of car, it will have a high activation level associated with the concept of 'car'; a less common member of the 'car' category (such as a luxury, for example), however, will likely have a lower level of activation. This notion of connectivity is essential to these models, as there are very specific ways in which units are connected. Specifically, the connections between units are what determine the pattern of activity that the entire framework processes. To clarify this point, each unit is said to have a 'fan in' and 'fan out' function (Rumelhart, 1989). The 'fan in' function is the amount of excitation or inhibition that is imposed on a given unit from other units connected to it. This function is a summed weight of all activity coming from other units. As for the 'fan out' function, this is the amount of activity that a given unit sends out to another unit. In addition, connectionist frameworks also contain 'hidden units' (Hanson & Burr, 1990), which are basically units that exist in the connections between regular units. These hidden units are hypothesized to act in a similar fashion as actual neural interneurons do--they assist in the transmission of activation levels between units by discriminating between necessary patterns and nonnecessary patterns, mapping only those activation levels that are needed from one unit to another. As a result, any knowledge that a connectionist framework has is derived from patterns of connected units, with each individual unit providing a fragment of the total knowledge base. This pattern of connectivity for any given concept is not a stable entity. Rather, any knowledge within such a system can be modified through experience (Rumelhart, 1989; Hanson & Burr, 1990). This occurs when connected units are continually active, causing the strength of the connection to increase. While this is occurring a previous connection that is no longer in use may lose strength. For example, if the only type car one was exposed to was a compact style, the connection car--size--compact would be highly active. However, if one then began to encounter other cars, that connection might lose some strength because a connection like car--size--luxury begins to alter the original pattern of connectivity.

In sum, while there are variations between different models (Hanson & Burr, 1990), any connectionist framework consists of processing units connected to one another, with knowledge (a) coming from the patterns of activation between the units and (b) being effected by the strength of various connections. As can be seen, connectionist KR systems are very complex, but they are also 'state of the art' in KR systems.

Representation. As mentioned above, the knowledge that a connectionist model possesses is contained within the connections between units, or more specifically, the weights that the connections have. This knowledge can be represented in several ways. First, any information can be completely local, which means that highly specific symbolic features are represented in a hierarchical fashion (e.g., my car--[red--midsize--Oldsmobile], similar to the semantic networks discussed earlier. The second type of representation is similar to the first, but the information is a bit more general (e.g., car--[Buick--Chrysler--Oldsmobile]. The third type of representation is more similar to connectionist networks as have been discussed here, where a distributed representation (McClelland & Rumelhart, 1985) connects all specific and general information in one network, with connections emanating from all units and extending to all units. The final representation is also a distributed network, like the third, but some of the nodes are completely blank--any type of information can be encoded in them for a given concept. This last type of arrangement is probably the most feasible way to handle the formation of new or abstract ideas.

Evaluation. Connectionist models of the mind are indeed highly advanced and integrated theories concerning knowledge representation. While some may criticize the complexity of the networks, one must also consider that the brain, which these systems are modelled after, is also exceedingly complex. A connectionist approach to the mind is also a superior way to view KR because 1) it deals with knowledge acquisition, something which the classical approaches to knowledge representation failed to do. On the other hand, although the existence of 'empty' units should enable connectionist networks to abstract structural information from stimuli and store such knowledge in an abstract form--a process which would facilitate the transfer of knowledge, these systems have difficulty with transfer tasks. This problem lies in the fact that connectionist representations tend to be fixed in the form in which input stimuli are formed, and such a knowledge store would seemingly interfere with the transfer process. With this in mind, it seems to be the case that connectionist networks need further development to rectify this very real problem if they hope to simulate actual human "thought".

Intermediate Summary

Thus far, this review has offered various perspectives on the nature of knowledge representation across many areas of cognitive psychology. With such a variety of theories postulated to handle the KR issue, it is easy to see that an analysis of the representational nature of mind, and with that the transfer process, is critical for understanding any subfield in cognitive psychology. The discussion will now review the issues surrounding knowledge representation and transfer within a field of psychological research that has existed for over twenty years but has recently ignited a minor controversy--implicit learning theory.

Implicit Learning: An Overview

For almost a quarter-century, the term implicit learning (IL) has been used to describe the process by which one acquires a deep understanding of a complex stimulus environment largely independent of conscious awareness of the specific elements of that environment (Reber, 1967). Such nonconscious learning has been found to occur across a variety of experimental tasks, with the factor common to all these procedures being that the stimuli presented to subjects is highly complex and rule-governed. This latter term refers to the fact that the stimuli in these experiments have a deep, meaningful structure embedded within them, in the same sense that a formal language grammar has a deep structure (Anderson, 1985). Due to the complexity of the tasks, subjects typically report little or no conscious awareness of this underlying structure, yet they are able to manipulate the stimuli in such a way as to facilitate performance during whatever procedure is being implemented.

A survey of the procedures used to investigate these nonconscious learning processes covers a variety of rule-governed tasks. To begin with, several researchers (Berry & Broadbent, 1988; Hayes & Broadbent, 1988) have investigated IL by utilizing an interactive production control task. This procedure involves providing subjects with a problem that has an initial and goal state, and the subjects task is to reach the goal state by engaging in some type of production-like behavior. To provide an example, subjects may work in a hypothetical sugar production factory where they are in charge of producing specified amounts of sugar. They accomplish this by manipulating the number of workers employed as well as considering the amount of sugar currently being produced. The amount of sugar produced by the chosen number of workers is determined by a formula, unknown to the subject, which considers the past amounts of sugar production and the current number of employees. Subjects are able to perform this task at above chance accuracy, but are not able to verbalize the formula. Another version of this task has subjects interacting with a 'computer person', with the object of the task being to get the 'computer person' to exhibit a certain type of emotion (Berry & Broadbent, 1988). The emotion that the 'computer person' shows is based upon both the subject's emotional response to the 'computer person' as well as past emotions of the computer person. As with the factory-like task, subjects participating in this procedure show little conscious knowledge of the underlying rule that determines the 'computer person's' reaction, yet they are able to control the computer's responses at above chance levels.

Another task used to investigate IL processes is known as a visual array search procedure (Lewicki, Hill, & Bizot, 1988). This task involves subjects attending to a visual array where certain stimuli appear in certain locations. After an initial exposure to various sequential spatial patterns, a stimulus may be presented in one location and the subjects task is to respond as quickly as possible as to where the stimulus will appear next. The locations are determined by some type of pattern which mandates which stimuli can appear at what times in which locations. For example, a computer monitor may be divided into 4 quadrants, and a stimulus (a number, a flash of light, a star, etc.) will appear in one of the four quadrants on one trial, then it will appear in either the same or different quadrant on the next trial, etc., until the entire pattern has been presented. In the Lewicki, et al., study (1988), the location of the first two stimuli was pseudo-random, in that they could not appear in the same location. The crucial locations were the next three positions, where the rule governing location was that the stimulus would appear in a quadrant that depended on where the stimulus appeared previously. For example, if on Trials 1 and 2 the stimulus moved vertically, then the stimulus would next appear in a quadrant horizontal to the previous location; if movement then occurred diagonally, the next location would be vertical, and so on. The subjects' task here was to indicate in which location the next stimulus would appear. As with the production task, subjects on this task were able to predict future locations at accuracy levels above chance, yet they reported only vague reasons for choosing the locations they did (e.g., "It seemed right"), and showed little conscious awareness of any structure to the patterns.

Recently, however, the idea that subjects performing such spatial tasks are abstracting complex rules from the patterns has been challenged. Perruchet, Gallego, & Savy (1990) argued that the knowledge acquired during such a task is not abstract at all. Rather, these authors feel that subjects simply store frequency information in regards to the spatial locations of stimuli, not complex patterns. In a reanalysis of Lewicki, et al.'s (1988) procedure and results, Perruchet and his colleagues found that the rule that Lewicki used to predict location was not the only rule that could capture the stimulus display. Specifically, Perruchet, et al. (1990) found that specific target movement patterns on Trials 1 and 2 occurred more infrequently than movement patterns from Trials 3 to 5, and by using frequency information, rather than knowledge of a complex rule system, subjects' predictions for future target locations were significantly faster for the more frequent patterns on Trials 3 to 5 as compared to predictions for Trials 1 to 2. Although Perruchet et al.'s argument does have some valid points (especially in regards to the issue of equally distributed stimulus locations--see Perruchet, et al. for a deeper discussion), it fails to completely contradict the findings to date on pattern learning, since subjects are still unable to provide valid reasons for their responses on these spatial tasks, i.e., implicit learning is still taking place.

While these latter two tasks have provided additional ways to investigate nonconscious learning processes, the oldest, most frequently employed procedure, and the one which has accumulated the largest data base on implicit learning, is known as artificial grammar (AG) learning (Reber, 1967, 1969; Brooks, 1978; Reber & Allen, 1978; Reber & Lewis, 1977; McAndrews & Moscovitch, 1985;

Mathews, Buss, Stanley, Blanchard-Fields, Cho, & Druhan, 1989; Perruchet & Pacteau, 1990). The stimuli used for this procedure are generated from a finite-state artificial grammar, which is basically a rule-generation system capable of producing stimuli which conform to certain structured patterns.

The procedure itself involves three phases. The first phase, referred to as the acquisition, or learning phase, has come to take several forms. In some cases, subjects simply attend to AG-generated stimuli (Perruchet & Pacteau, 1990), while in other studies one must attend to and memorize presented stimuli up to a specified performance criterion (Reber, 1969; Reber & Lewis, 1977). Once this phase is completed, subjects are informed that the sequences they saw followed rules which determined the order of stimuli within each sequence, although no mention is made as to what the rules actually are. Subjects are then shown additional stimuli, with some of them being permissible sequences and others containing a violation in a certain position (e.g., from the grammar in Figure 1, the sequence MHXH is permissible, but HHXH is not). The subjects' task is to decide if each presented stimulus is permissible or not. Once this phase has been completed, subjects are interviewed in order to ascertain what overt, declarative knowledge, if any, of the grammar the subject was able to acquire from the two previous phases. The typical result from experiments of this nature is that subjects make their permissibility decisions well above chance levels, yet fail to verbalize concrete rules of the grammar that they have been using.

Recently, researchers have developed yet another technique that combines both pattern learning and AG learning in the same task. In this technique, subjects are presented with either visual (Cleeremans & McClelland, 1991) or auditory (Manza, 1991) patterns, with the rules that determine the location and/or order of stimuli determined by artificial grammars. The results of the application of this new technique is virtually identical to the results of the 'traditional' pattern learning or AG learning tasks: subjects' performance exceeds chance levels, yet they report little, if any, conscious awareness of the underlying structure of the patterns.

Perspectives on the Representation of Tacit Knowledge

Researchers within the IL field have taken results from such artificial grammar learning tasks and applied them to theories concerning how nonconscious information is represented. The results that have shed the most light on the representational issue are those that have come from transfer studies, where the physical characteristics of information presented during learning (e.g., an artificial grammar item such as MHHXH) and testing (e.g., LJJZJ) phases differ, yet they share an underlying structural commonality. The power of the transfer design comes from the manner in which the form(s) of the information presented during learning and testing may be manipulated independent of each other.

By interpreting transfer results, some researchers have argued for abstract representations of tacit knowledge (Reber, 1989; Mathews, 1991), while others have countered with an instantiated account of nonconscious knowledge representation (Brooks, 1978; Brooks & Vokey, 1991). A third position, known as the fragmentary view, is not based on transfer findings, yet is relevant to the current discussion on the representation of implicitly acquired information (Dulany, Carlson, & Dewey, 1984; Perruchet & Pacteau, 1990, 1991). These opposing positions have led to a debate among IL researchers as of late (Perruchet & Pacteau, 1990, 1991; Mathews, 1990, 1991; Reber, 1990), but in order to understand the current debate it is necessary to review each of these alternative representational hypotheses to determine which position if any, is most accurate.

The Abstractive View

Overview. The oldest perspective concerning the nature of a tacit knowledge representation is known as the abstractive view (Berry & Broadbent, 1988; Lewicki, Czyzewska, & Hoffman, 1987; McAndrews & Moscovitch, 1985; Mathews, et al., 1989; Reber, 1969, 1989, 1990; Reber & Lewis, 1977; Reber & Allen, 1978). The term abstract, as used here, suggests that the initial form of stimuli is modified by a 'taking away' of certain elements of the initial stimulus. As a result of this abstraction process, the new item stands apart, as a separate entity, in comparison to the initial instantiation of the stimulus. In the case of implicit knowledge, the superficial surface elements of artificial grammar-generated items are removed from the initial representation, leaving a new stimulus which is a modification of the original item. This new representation contains little, if any, information pertaining to specific stimulus features; the emphasis here is on structural elements within stimuli. For example, if the letter string MKKP was presented to a subject, the abstractive view would propose that the individual letters making up the item could be abstracted away, leaving a representation of the form 'One occurrence of symbol type 1, followed by two occurrences of symbol type 2, ending with an occurrence of symbol type 3'. Several transfer studies have provided some support for this abstractive theory of tacit knowledge representation.

The initial IL transfer study was conducted by Reber (1969). In this experiment, subjects were run through two learning phases of an artificial grammar task. Initially, all subjects were presented with letter strings in groups of three, with the criterion for each triplet being to correctly reproduce each member of the triplet on the same trial. The critical manipulation came on the second learning task, where four different types of transfer tasks were employed. Although the task itself was identical to the initial learning phase, the transfer manipulation varied between groups. One group of subjects were presented with items that differed from initial learning items in their orthographic form, although the grammar was identical; another group had the symbol sets remain constant between phases, but the rule structure changed; a third group had their symbols and grammar changed between phases, and a final group served as a control, with letter set and grammar remaining constant between phases. The significant results, from the second learning task, were that while changing the grammar resulted in performance levels that were significantly lower than control performance, changing the letters set only did not result in a performance decrement, as compared to control subjects. This lack of a performance decrement as a result of the transfer manipulation led Reber (1969) to conclude that implicit knowledge is most likely represented in an abstract, stimulus-independent form. Building upon this assumption and applying a transfer design across sensory modalities, Howard & Ballas (1982) were also able to obtain positive transfer within the framework of an IL procedure. These investigators presented subjects with environmental sounds (e.g., squeak, drip, hiss, clang, & flush) in either a visual (the written words) or auditory (the sounds represented by the words) manner during an observation-style learning phase of an AG task. During the subsequent testing phase, one group of subjects, which had received visual stimuli during learning, was now presented with auditory stimuli. The performance of this group on the well-formedness task did not differ significantly from subjects who were presented with stimuli from the same sensory modality (visual or auditory) during the learning and testing phases. The authors concluded that Reber's abstractive view was supported by their data.

Finally, Mathews, et al. (1989) employed a transfer design investigating the effects of changed letter sets over a four week period. In this study, each subject participated in four sessions, each one week apart. For each session, subjects initially studied a list of items generated from an artificial grammar. They were then tested by a multiple choice-type task: successive sets of five items were presented, with each set consisting of four nongrammatical items and one grammatical item; the task was to select which of the five stimuli was most grammatical. The stimuli for each individual learning/testing session utilized the same letter set during the two phases of the experiment. The transfer element entered into the design in each successive week, in that a different letter set, with the same deep structure as previous sets, was used to instantiate the grammar for each session. For example, during week 1, a subject might have encountered the item DKKDS during testing, and during the second week that item might have been presented as WMMWZ. Results indicated that subjects who received stimuli from different letter sets each week (the transfer subjects) were able to perform the multiple-choice task significantly above chance, although their performance was significantly lower than those subjects who were presented with items from the same letter set each week (control subjects).

Evaluation. Taken together, the results of these studies suggest that there is a strong possibility that tacit knowledge is represented abstractly. The fact that subjects can accurately discriminate between grammatical and nongrammatical items that are instantiated with different letters than those originally learned suggests that subjects in implicit learning experiments induce structural features of stimuli rather than surface features. However, this abstractive view of the transfer findings has been challenged by alternative explanations, outlined below, that must be considered in order to capture the full essence of the issues surrounding implicit knowledge representation.

The Distributive View

Overview. This perspective, championed by Lee Brooks and his colleagues (Brooks, 1978; Brooks & Vokey, 1991) proposes that tacit knowledge is represented not in an abstract manner, but rather in an exemplar-based fashion. That is, when a stimulus is encountered, regularities are not extracted and represented in a prototypical fashion. Rather, Brooks proposes that individual items are stored into memory in their original form, with the resulting knowledge base filled with numerous instances of specific items. Taking the item MKKP once again, the distributive view would represent this item in its raw form, MKKP. This perspective of nonconscious processing is based on a seminal study (Brooks, 1978) aimed at assessing whether the implicit learning process occurred by way of abstraction or instantiation. One group of subjects was exposed to 30 paired associates (PA-1), with one member of each pair being a string of letters generated from an artificial grammar and the other member being either the name of an animal or the name of a city (e.g., VTRR--tiger, VVTRVT--New York). Half of the letter strings for these PA-1 items were generated from a certain artificial grammar (Grammar A), while the other half were generated from a different grammar (Grammar B). The distribution of animals and cities was counterbalanced between items from the two grammars, and the letters used to instantiate both grammars were identical. A second group of subjects (PA-2) was also exposed to 30 paired-associates, but instead of pairing letter strings with animal and city names, 15 items from Grammar A were paired with the word "city" and 15 items from Grammar B were paired with the word "animal" (e.g., VTRR--animal, VVTRVT--city). For both groups, once the pairs were memorized to a criterion, subjects had to complete a sorting task, where they had to sort 30 additional letter strings into one of three categories: Old World, New World, or Neither. These three categories corresponded to Grammar A, Grammar B, and neither grammar, respectively. An additional group of subjects, referred to as concept learners (CL), were not initially exposed to paired-associate learning. Instead, these subjects were presented with the same 30 items that groups PA-1 and PA-2 had to sort, and were instructed to classify the strings into either Grammar A, Grammar B, or Neither. These subjects were also informed as to the accuracy of each individual classification. The resulting data indicated that 1) there were no differences in the sorting accuracy of the two PA groups, and 2) all three groups (PA-1, PA-2, and CL) sorted test items into the three categories well above chance levels, although CL subjects performed significantly less accurate (Mean=46% correct) than the combined score of the two PA groups (Mean=62% correct). In interpreting this higher sorting performance of the two PA groups as compared to the CL group, Brooks concluded that when subjects memorize and store individual instances of items, they develop a more accurate representation of a stimulus environment. This position is clearly different from the abstractive view, although a study by Reber & Allen (1978) shed some light onto why the PA subjects in Brooks' experiment performed better than the CL subjects.

Reber & Allen (1978) explained Brooks' results as possibly being due to the nature of the learning task, in that "...the PA task itself...encourages individuated memorial representation and reasoning by analogy quite independent of the subterfuge of Brooks' procedure" (p.196). To test this assumption, they utilized two different grammars (with identical letters) and two types of learning phases (paired associate (PA) and observation (OBS)) in a within-subjects design. For each subject, if Grammar A was used for PA learning, Grammar B would be used for OBS learning, and vice-versa. The PA task involved the presentation of 20 letters strings, each paired with the name of a city, with a performance criterion established for each item. For the OBS task, on the other hand, subjects were presented letter strings and were instructed to simply pay attention to them. Immediately following each task, subjects were presented with a classification task, where they had to classify each of 100 letter strings as either (1) being a possible sequence for a city (following PA learning) or (2) being well-formed according to the rules of the grammar that the strings were generated from (following OBS learning). Half of the subjects went through the PA-classification task first and half went through the OBS-classification task first. On the average, 19.8 days separated the two procedures. Reber & Allen (1978) found that OBS subjects classified test items significantly more accurately than PA subjects (81% vs. 74%, respectively), suggesting that subjects seem to have an instantiated representation following the PA procedure and an abstract representation following OBS learning.

Another study has recently looked at this effect of the demand characteristics of the acquisition portion of IL procedures. Manza (1992) had two groups of subjects each study 2 list of 20 items each by utilizing an observation method similar to that employed by Reber & Allen (1978). The items on the two lists for Group 1 subjects were identical, but Group 2 subjects studied items in different surface instantiations. For example, the item TXXS appeared on both study lists for Group 1 subjects, while Group 2 subjects were instead presented with TXXS and its structural 'twin', MKKP. Once both lists were studied, all subjects classified Control items (stimuli instantiated in the same letter sets as the items that were initially studied) and Transfer items (stimuli that shared structural, but not surface, elements with Study List items) as conforming or not conforming to the rules of an artificial grammar. Results indicated that by encouraging, general, transfer-appropriate processing (Graf & Ryan, 1990), one's representation could be either exemplar-based or abstract, a result supporting the suggestion made earlier by Reber & Allen (1978). Specifically, with both groups performing significantly above chance, the performance of Group 2 subjects in classifying Transfer and Control items did not differ from each other (both Means=57% correct), while Group 1 subjects showed significantly higher accuracy with Control items as compared to Transfer items (Means=63% vs. 53% correct, respectively). This result suggests that when one is led to study information in a conceptual manner, the ensuing representation may be primarily abstract, but when one focuses their attention on individual instances, the resulting representation may be primarily exemplar-based.

Since transfer studies, like the ones described above, can provide a wealth of information regarding the representation of implicitly acquired knowledge, supporters of the distributive view (Brooks & Vokey, 1991) have offered yet another explanation of how tacit knowledge is represented in reference to the transfer procedure. Brooks & Vokey (1991) contend that transfer performance can be accounted for by subjects making analogies from presented stimuli to similar, stored examples during the classification task. According to this perspective, in dealing with transfer within a letter set, if a subject had a stored instance of, for example, the letter string PTTSXS, and was presented the string PHXSPS, a subject might decide that the latter string is grammatical because of the similar P at the beginning and alternating S's near the end. In terms of transfer between different letter sets, Brooks & Vokey (1991) argue that nonconscious transfer is accomplished by a subject drawing an 'abstract analogy' between two items, based on similarity. For example, if one studied the item PTTXS and later had to classify BMMZQ, one would reach their decision by noticing that, for example, "P is replaced by B, T replaced by M, X by Z, and S by Q." If it were found that this type of analogy were truly being formed, it would weaken the abstractive view. While there do exist data in support of the distributive view, there is no concrete data to support the type of mapping that Brooks and Vokey describe. However, the Manza (1992) study described above tested for the existence of these analogies, and the results of this investigation showed that subjects could not determine, above chance, which letters from one letter set were replaced by which other letters from a different letter set. However, this is just one finding, and by no means conclusive. Therefore, it appears to be the case that both the distributive and abstractive perspectives on implicit knowledge representation offer plausible explanations for the questions surrounding the representation of complex information.

Evaluation. In reviewing the distributive model of tacit knowledge representation, a picture begins to emerge as to the form in which tacit knowledge is stored. An exemplar-based memorial system differs from an abstract system, but it is most likely the case that both forms of knowledge may be cognitively processed. If anything has been accomplished by the above review, it serves the purpose of pointing out that abstractive and distributive mental components may have a synergistic effect on one another.

The Fragmentary View

Overview. The final position in the current debate focusing on the nature of nonconscious knowledge representation is known as the fragmentary perspective (Dulany, Carlson, & Dewey, 1984; Perruchet & Pacteau, 1990, 1991). This view differing radically from the abstractive perspective and mildly from the distributive view, holds that the form of representation that results from IL procedures is a fragmentary one, consisting of (1) specific knowledge of letter bi- and tri-grams (e.g., the string TSSXPH is represented as TS, SX, and PH, not as an integrated whole), as well as (2) salient knowledge of frequencies of covariational patterns. Independent of the actual nature of the stimulus display, supporters of this position suggest that subjects 'split' a full stimulus into several fragmentary pieces, chunks of data which stand alone in a distribution of many different, but similar, fragmentary units.

The first explanation of, and the methods used to produce, the data that claim to support this theory was gathered by Dulany, et al. (1984). Their experiment had subjects initially partake in a typical AG learning phase. For the testing phase each subject was presented a list of grammatical and nongrammatical items on a sheet of paper, with their task being to: (a) identify which strings were nongrammatical by crossing out the section of such strings that they thought violated the rules of the grammar, and (b) identify grammatical strings by underlining components of such strings that they believed were significant in making them grammatical. The authors claimed that the subjects' performance on this task, which was above chance (and paralleled previous artificial grammar learning findings (e.g., Reber & Lewis, 1977, Reber & Allen, 1978), where more traditional Yes/No well-formedness tasks were utilized during testing), reflected their conscious knowledge of the rule structure of the grammar. It was further argued that the knowledge base acquired during the procedure consisted primarily of specific bi- and tri-gram patterns. In a comment on Dulany et al.'s evaluation of the form of the tacit knowledge base that subjects acquire during the AG task, Reber, Allen, & Regan (1985) countered with the argument that the task that Dulany et al (1984) implemented to assess their subjects' knowledge of the grammar did not tap explicit processes, as Dulany and his colleagues claimed. Rather, it was proposed that the recognition task that was used may not have been a valid measure of explicit knowledge, a claim which has subsequently been supported by research in implicit memory which has utilized recognition procedures as measures of nonconscious knowledge (Graf, Shimamura, & Squire, 1985; Roediger & Blaxton, 1987; Schacter & Graf, 1989). In addition, Reber et al. (1985) posited that even though subjects might acquire some fragmentary knowledge of the rule structure of the grammar during an AG task, their knowledge, for the most part, is an abstract understanding of deep covariational patterns in the stimuli. One final point to Reber et al.'s argument is that Dulany and his colleagues failed to distinguish between consciousness during the AG task and the capacity for retrospective, conscious, analysis later on--a point which is crucial to understanding nonconscious earning processes. After this brief skirmish, the fragmentary view seemed to fall out of sight for a few years, until it was recently reborn by two French researchers, Pierre Perruchet and Chantal Pacteau.

Perruchet and Pacteau (1990), in a series of experiments, claimed to have evidence that the knowledge that one acquires during an AG task is primarily fragmentary, consisting (like Dulany, et al) of patterns of bigrams. The authors reached this conclusion in Experiment 1 of their 1990 paper by utilizing the following procedure. During learning, letter sequences generated from an artificial grammar were listed on a sheet of paper in either one column in their whole form (e.g., TXXS) or in three columns of 17 bigrams each--a fragmentary form. The frequency of bigrams was controlled to match that of the occurrences in the "whole form" items. After studying either list, all subjects were shown whole form items and had to make grammaticality judgments concerning these items.

Subjects who had whole form learning were correct 61.5% of the time in the grammaticality decisions, while those subjects who studies bigrams during learning had a percent correct score of 57.2, a score which was significantly above chance, yet significantly lower than the performance of the whole form group. Despite this obvious discrepancy, Perruchet and Pacteau (1990) claimed that this above chance performance of the bigram-learning group is evidence that tacit knowledge is primarily represented in a fragmentary form. However, since the whole form subjects performed significantly better than the bigram group, Perruchet & Pacteau's conclusion seems overstated. Another study in the same investigation (Experiment 3) had subjects presented with a list of whole form strings during learning, followed by a list of bigrams during the testing phase. Subjects' task here, however, was to rate each bigram in terms of how confident they were that each item was part of a whole string. To do this, subjects used a 6 point scale, ranging from 1 (meaning 'sure that this pair was not part of a whole string') to 6 (meaning 'sure that this pair was part of a whole string'). The mean rating score on this task was 3.31, and the authors claimed that this result provided additional support to their fragmentary perspective. However, scores between 3 and 4 on this scale are indications of uncertainty about the status of each bigram, so it is not clear that subjects are entirely confident in judging the well-formedness of rule-governed bigrams. Nevertheless, the authors maintain their view that synthetic grammar learning results in the formation of a knowledge base that is fragmentary in form.

Evaluation. Noting these inadequacies of Perruchet & Pacteau's (1990) conclusions, Reber (1990) and Mathews (1990, 1991) attempted to show that Perruchet and Pacteau are erroneous in claiming that all tacit knowledge is represented in a fragmentary fashion. Both pointed out that although some tacit knowledge may be represented in a fragmentary form, there are other ways to represent nonconscious information. Mathews (1990, 1991) goes on to point out that transfer studies, which are believed to be the best evaluators of the form of a knowledge base, have not been attempted to support the fragmentary perspective, and the transfer results that have been reported in the implicit learning literature cannot be accounted for by a fragmentary argument. Perruchet and Pacteau (1991) countered this latter suggestion by taking a stance that resembles the distributive account of the transfer process. Specifically, they postulated that subjects participating in a transfer task try to retrieve specific, stored exemplars that have similar patterns to the transfer items. This new position taken by Perruchet and Pacteau (1991) seems to be drifting them away from their fragmentary perspective, towards a more distributive view, a point also mentioned by Mathews (1991). This failure to sufficiently account for transfer, as well as poor performance by subjects who must determine the grammatical status of isolated bigrams, are the main reasons why there is reason to believe that the fragmentary view is not the only possible perspective pertaining to the representation of nonconscious knowledge. This is not to say that Perruchet and Pacteau are totally wrong--quite the contrary. It is most likely the case that some tacit knowledge is represented in a fragmentary form, but most of it isn't.

Summary

From this review, the characterization of a tacit knowledge base seems to be one in which information is (1) acquired and held primarily outside of conscious awareness, and (2) represented in abstract, distributed, and fragmentary forms. However, as with knowledge in general, knowing solely what the content of a tacit knowledge base is neglects another area of implicit learning. This area has to do with the mechanisms responsible for the structuring and processing of tacit knowledge. In an attempt to determine how a tacit knowledge base is processed and structured, several researchers have recently proposed different models of nonconscious processing. Therefore, to get a complete account of the existing knowledge base on implicit representations, the discussion will now review the various models proposed to account for the structuring of tacit knowledge within the mind.

Models of nonconscious functioning

THIYOS

Recently, Roussel, Mathews, & Druhan (1990) proposed a computational model of implicit learning that is predicated upon a production-based classifier system (Holland, et al., 1986). This model, termed THIYOS, is basically made up of condition-action pairs of the form "IF [a certain condition is met] THEN [perform a certain operation]." Such production rules are the heart and soul of production system models of the mind (Anderson, 1983). THIYOS functions by way of two different algorithms, referred to as a genetic algorithm and a forgetting algorithm. The genetic algorithm is used to create new rules and/or alter existing ones. This is accomplished by taking a rule and splitting it, and then applying each half to create a new rule. The forgetting algorithm, on the other hand, retains one part of a stimulus and discards, or 'forgets', the rest of the stimulus in an attempt to develop new rules in a more economical fashion. An example will help to clarify these processes. Let's say that two stimuli, VVTRDV and TVTRDV, are presented to this system. The genetic algorithm will first make a copy of each sequence and then split the first one into VVT and RDV and the second into TVT and RDV. It will then combine different halves in attempt to create new rules (e.g. VVTTVT and RDVRDV), with these new rules being stored with the presented ones. The forgetting algorithm, on the other hand, would take a sequence, e.g., VVTRDV, and remove the VVT part to create the rule "IF a sequence ends in RDV, THEN it is a valid rule." As additional sequences are presented to the system, they are checked against the existing rules in order to determine if it should be chosen as a valid rule, an invalid rule, or reformed to create a new rule. If a presented sequence finds a match in storage, then it is chosen as a valid rule. If a presented sequence is found to contain a violation, it is chosen as an invalid rule, and if no match at all is found, a new rule will be created, depending on which algorithm is functioning. Thus, the THIYOS model postulates that both specific and general knowledge can be stored in a production system that basically contains numerous condition-action rules that determine how stimuli are formed and processed. While the inclusion of both general and item-based information is a significant feature of this model, THIYOS would most likely have a difficult time handling transfer, a point which renders this model incomplete in terms of fully characterizing the implicit learning process. This model, for example, has no mechanism for converting one symbol set to another (e.g., understanding that CCT and HHQ are structurally equivalent), so while it can model the acquistion of tacit knowledge somewhat, it needs to be modified in some way in order to handle the transfer of tacit knowledge, a process that has been shown to occur empirically (Reber, 1969; Howard & Ballas, 1982; Mathews, et al., 1989).

Competitive Chunking

Another proposed model of implicit learning was recently formulated by Servan-Schreiber and Anderson (1990). While this model seems to capture the data from the implicit learning process, it falls short of completely representing the acquisition of nonconscious information due to the fact that it too cannot adequately deal with transfer. This approach, known as competitive chunking (CC), proposes that the main processor of information acquired during artificial grammar learning is a chunking mechanism that represents tacit knowledge in the form of chunks. The chunks are postulated to be organized in long term memory in a hierarchical fashion, similar to various semantic-network models (Chang, 1986). The main CC processor holds in its memory two bits of information about every chunk in the system: (1) the strength of the chunk, i.e., how often it has been used in the past, and (2) what a particular chunk's immediate subchunks are. Processing in this system is said to occur in a bottom-up fashion, proceeding from the most basic to the most complex chunks. For example, the representation of the string TTXCZCVT might be as follows:

1. [TTXCZCVT]
2. [(TTXC) (ZCVT)]
3. [(TT) (XC) (ZC) (VT)]
4. [(X) (T) (X) (C) (Z) (C) (V) (T)]

Processing would initially occur from the bottom (#4) chunks up to the complete, eight unit chunk at the top of the hierarchy (#1). Each time the TTXCZCVT stimulus is encountered, it requires less processing than on its previous presentation, with processing starting at higher levels in the hierarchy on each successive presentation. For example, the first time this item is encountered, it might be processed from Level 4 through Level 1. On the next presentation, if the delay between the first and second presentations is not too long, the item might require processing only from Level 2 to Level 1, due to the increased strength of the stimulus as compared to when it was first encountered. Eventually, processing only requires accessing the top level of the representation. As a result of this type of framework, CC holds that specific instances are stored, much like the distributive view of tacit knowledge. Working under the assumption that if one simply stores instances, with no reference to general structure, competitive chunking could not transfer tacit knowledge between stimulus domains, rendering this model, like THIYOS, incomplete in terms of capturing the implicit learning process.

Parallel Distributed Processing: SRN

The final model of IL currently available is known as a simple recurrent network (SRN). Cleeremans & McClelland (1991), who have formulated the model, propose that implicit learning can be modelled by implementing a three level representational structure in the following manner. Level 1 contains the current input into the system, level 2 consists of hidden units that encode relevant features of a stimulus and level 3 contains the output of the system. In addition, connected to the input unit is a context unit that stores the entire stimulus content at time t. After a stimulus is processed at the hidden unit, it is 'fed back' to the context unit, where the next stimulus is encountered. For example, if the letter set TSXS was presented to the system, the letter T is initially processed, and then sent to the hidden unit. At this point, a copy of the T is back propagated to the context unit and stored there. Then the next stimulus, S, is input into the hidden unit, where it is processed and a copy of it is sent to the context unit, where it becomes spatially associated with the previous T. This process continues until there is no stimulus present. At this point, the entire string, TSXS, is sent to the output unit and is stored in memory. As new stimuli are input, they are compared to stored stimuli, in the context unit. The main storage component of this system, therefore, lies in the context unit, where sequences are stored and, over time, gain strength as they are continually encountered. This system, like the previous two models, seems to store data in a stimulus-specific fashion, and applying the main critique of THIYOS and CC to the simple recurrent network model, transfer appears to be a difficult process for the SRN to represent.

Summary

Each of these models presents viable accounts of the framework and processing mechanisms behind the acquistion and utilization of tacit knowledge. Specifically, they all provide an adequate account of the retrieval processes involved in a nonconscious representational system. However, where each of these models fails is that none of them can adequately handle the encoding of information in a fashion that is independent of the form in which data is originally input into a system, a process which would facilitate the transfer process. While one of them (THIYOS) does seem quite capable of developing abstract representations (by way of the forgetting algorithm), the representations are still limited to a specific stimulus domain. This lack of ability to develop domain-free representations is a major shortcoming of these models, and the end result of the current investigation will hopefully provide insights into how the transfer process can be handled by some model of the implicit learning process.

Intermediate Summary

This review of the current state of the literature on nonconscious knowledge representations has served the purpose of drawing attention to the data that exist in the implicit learning field concerning the representation of complex information. The models that exist all have flaws, and the theories proposed to account for the content of nonconscious thought (abstractive, distributive, and fragmentary) do not have an overwhelming amount of support in favor of any one over the others. While transfer tasks have been carried out in some cases, the results have not been conclusive. Therefore, to explore further the issues surrounding the representation of tacit knowledge, four implicit learning artificial grammar learning transfer experiments were conducted. The first two studies investigate transfer effects across different physical stimulus forms (letter sets) by implementing both a between (Experiment 1) and a within (Experiment 2) subjects design. Experiment 3 probes transfer effects across different sensory modalities, while the final study looks at transfer across both physical form and sensory modality.

As there is a considerable amount of overlap between the procedures employed in the four studies, a general method section will outline the basic method used. The only information, other than that in the General Method below, that will be discussed in each method section will be methods and procedures that were specific to the particular experiments.

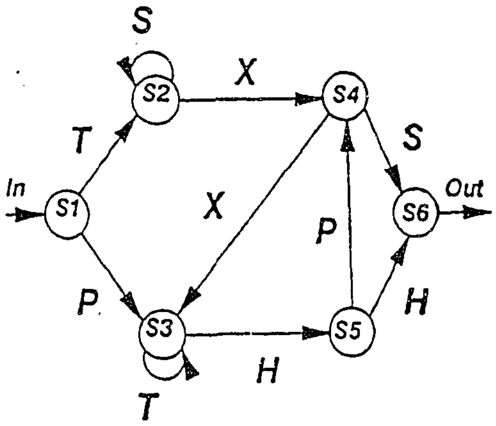
General Method

Subjects and Materials

Brooklyn College undergraduates that were either monetarily compensated or received course credit served as subjects in all four studies.

Stimuli for the experiments were generated from various finite state artificial grammars, similar to the one in Figure 1 below. Although these systems can generate an infinite number of stimuli of infinite lengths, only selected items that best represented the structure and item-distribution of the grammars were selected for presentation. Both Grammatical (G) and Nongrammatical (NG) stimuli, ranging in length from 3 to 8 symbols, were generated for presentation. The former items are formed by simply following any permissible path in the grammar and 'picking up' stimuli that appear along the path. Nongrammatical items, on the other hand, are generated by replacing a symbol at some position in the grammar with a symbol that does not conform to the structure of the grammar. For example, the grammar in Figure 1 can generate the item TSSXS by following a certain path; this item is considered Grammatical.

However, if the initial T is replaced by the letter H, the new item HSSXS is not permissible; this item is considered Nongrammatical.



Grammatical items	Nongrammatical items
PHH	TXH
PHPS	TXPH
TSSXS	XXSHT
TXXTHH	PTHHHH
PHPXHH	HSTXHHS
TSXXTHPS	PHXPHXPX

Figure 1. A finite state artificial grammar, including examples of Grammatical and Nongrammatical items.

Procedure

Learning. Subjects would initially be instructed that they were to basically memorize presented items up to a set criterion; no mention was made as to the rule-governed nature of the stimuli or that any task would follow the learning procedure. Only Grammatical items were presented during learning, but the subjects were not made aware of this fact until they had completed this portion of the procedure. Except for two conditions in Experiment 3 where stimuli were read aloud to subjects, all learning items were presented in the center of a computer monitor for 5 seconds, after which the stimulus was removed from the display and the subject had to reproduce that item. Reproduction was accomplished by pressing the keys on a keyboard that corresponded to the presented items. This sequence of presentation and reproduction continued until a criterion was met, after which the next item would be presented. Once subjects had met the performance criterion for all learning stimuli, they were allowed to relax for 3 minutes, after which the instructions for the next task were administered.

Testing. Upon completion of the acquisition task, subjects were informed for the first time that the items they studied previously were rule-governed in the sense that there were a set of rules which determined the order of the symbols within each item. No mention was made as to what the specific rules were. Subjects were then instructed that their task was to decide if each of a series of presented stimuli was well-formed according to the rules of the grammar. An equal number of Grammatical and Nongrammatical items were presented to each subject. Each item appeared in the center of a computer screen until the subject decided 'Yes' (meaning 'yes, this item is well-formed') or 'No' (meaning 'no, this item is not well formed'). Decisions were made by depressing the 'Y' or 'N' key on the keyboard. Once the decision was indicated, the stimulus would be removed from the display and a confidence rating scale would appear before the subject. Subjects' task here was to rate their confidence in the accuracy of their previous decision, according to a scale ranging from 1 (not confident at all) to 7 (very confident), with 4 being neutral. Once a rating was made by depressing the appropriate number key on the keyboard, the scale would be removed and the next item would be presented.

Interview. Once a decision and confidence rating was made for the last stimulus, subjects were asked a series of questions (shown in Appendix A) designed to determine: (a) what knowledge about the grammar was acquired from the two phases of the experiment, and (b) what types of mental activity were being engaged throughout the two phases of the experiment. This interview was conducted over several stages. Subjects were first handed the questionnaire and asked to answer all of the questions as completely as possible. This process took approximately 5-10 minutes to complete. Once subjects indicated that they had completed the questionnaire, the experimenter would attempt to get the subject to recall any additonal rules they may have been using by asking them repeatedly if they could recall any additional information. This would go on for approximately another 10-15 minutes, with the total time of the interview, therefore, lasting 15-25 minutes.

Experiment 1

As detailed above, several implicit learning studies have looked at transfer across different orthographic forms. However, these studies (Reber, 1969; Mathews, et al., 1989) did not change the form of the stimuli between the learning and testing phases. Reber (1969) employed two consecutive learning tasks, while Mathews, et al. (1989) changed the letter sets each week, but not during each session. Recently, Brooks and Vokey (1991) altered the letter set between the learning and testing phases, but their result were inconclusive in terms of the issues surrounding the representation of tacit knowledge. Therefore, to determine how implicitly acquired information is represented, and to differentiate the effects of altering the surface forms of stimuli immediately after initial acquisition, Experiment 1 required subjects to use knowledge acquired while working with stimuli with one physical instantiation to make decisions about the structural integrity of stimuli instantiated with different physical forms. If subjects are actually utilizing the structure of the grammar during the transfer task, they should be able to perform the well-formedness task at levels above chance, and the individual letter sets used during both learning and testing should have no effect on performance, independent of the transfer manipulation. The results of this study will hopefully shed more light on the issue of the representation of tacit knowledge.

Method

Subjects, Design and Materials

Ninety Brooklyn College students were solicited to participate in this study. Subjects were compensated either by being paid a flat fee of $5.00 or receiving course credit for their participation. A 2 (training letter set A or B) X 3 (testing letter set A, B, or Scrambled) between subjects design yielded six experimental groups (when described below, the first letter of a group's abbreviation refers to the letter set presented during learning, while the second letter refers to the testing letter set).

The artificial grammar in Figure 2 was used to generate the stimuli for this study; all learning and testing stimuli are displayed in Appendix B. As can be seen from the figure, each symbol has a direct correlate at each position in each letter set. To illustrate this correspondence, the grammar in Figure 2 can generate the string MXXMMQBH using letter set A. By mapping the letters directly from one set to another, this item can appear as LZZLRWJ in set B, QHHQQXMB in set A*, and RJJRRZLW in set B*.

The A-A and B-B groups served as controls, as subjects in these groups were presented with stimuli from the same letter sets during the two main phases of the experiment. The A-B and B-A groups were presented with items that conformed to the same grammar, yet were instantiated with a different letter set during the two phases of the experiment. Finally, to see if the individual letters provide interference in developing a representation, the stimuli

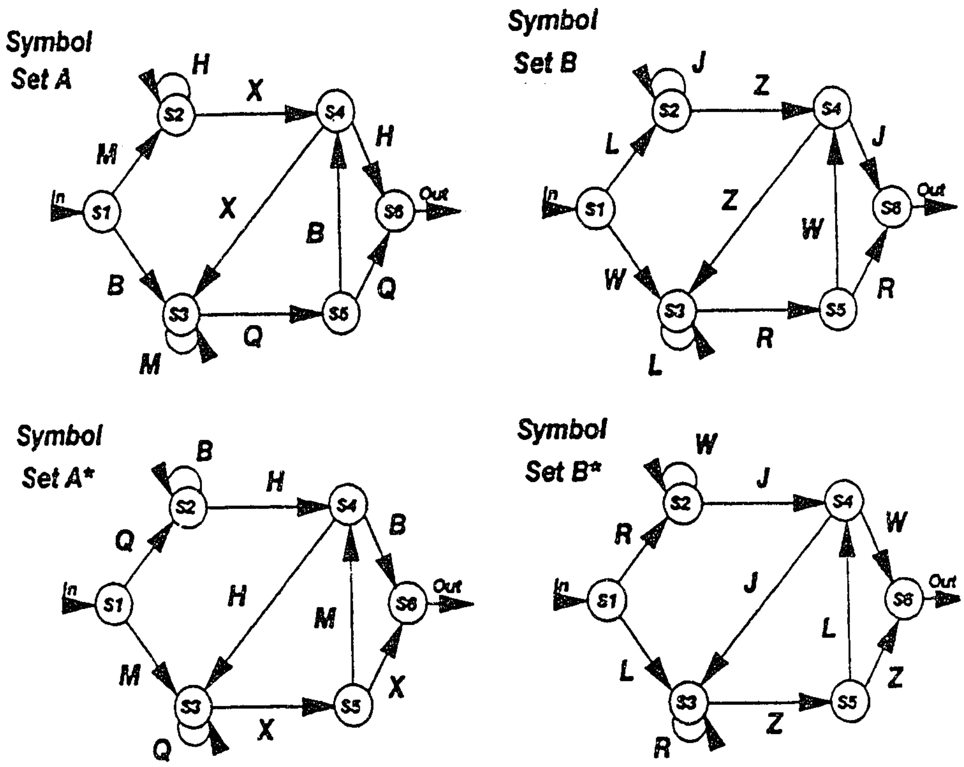


Figure 2. The artificial grammar used to generate stimuli for Experiment 1, with its four instantiations.

presented to groups A-A* and B-B* consisted of the same letters during learning and testing, but the positions of the letters within the grammar were switched for items used during the testing phase, as compared to the learning phase.

Procedure

Learning. During this phase, subjects were initially informed that their task was simply to memorize and reproduce each of 20 individually presented items that would appear at the center of a computer monitor. These stimuli were presented in the same order for all subjects, with the criterion for each item being two consecutive correct reproductions.

Testing. After being instructed about the general nature of the rules of the grammar, Regular Transfer subjects (conditions A-B and B-A) received the additional information that although the letters in each test item would be different, the rules that determine the order of the letters within each item are identical to the rules that the determined the orders of the items studied during learning. Subjects in the Scrambled Transfer conditions (A-A* and B-B*) were informed that the test items will consist of the same letters as in the learning items, but their positions within each stimulus will be different.

As for the task itself, subjects judged the well-formedness of 25 Grammatical and 25 Nongrammatical items, individually presented in a pseudorandom order--G and NG strings were mixed within the presentation list, but all subjects received this mixed list in the same order. Once the set of 50 was completed, presentation continued with the first item on the list, concluding again with the 50th. Therefore, the testing phase consisted of 100 trials (50 strings X 2 presentations each).

Interview. At the conclusion of the testing phase, subjects were administered the questionnaire shown in Appendix A, which was designed to assess what knowledge, if any, subjects could verbalize about the rule structure of the grammar.

Results and Discussion

Learning Phase

Trials to Criterion

In order to determine if the individual letter set used to instantiate the artificial grammar had any effect on learning the items presented during the initial study phase, a one-way ANOVA was conducted on the trials to criterion data. This analysis confirmed the hypothesis that the letter set in which stimuli are presented has no significant effect ($F < 1$) on learning, as the mean trials to criterion for each letter set was virtually identical (Letter Set A mean=2.9 trials, Letter Set B mean=2.8 trials).

Testing Phase

Percent Correct

Comparisons Between Control, 'Regular', & 'Scrambled' Transfer. Analyzing percent correct scores allows a general insight into the occurrence (or lack of) positive transfer. Initially, a one-way ANOVA revealed a significant difference between the six groups in terms of their overall accuracy in making well-formedness decisions, $F(5,84)=10.7$, $MSe=112.7$, $p < .0001$. To determine exactly where this difference occurred, several planned comparisons were carried out and found no significant differences between 'Regular' Transfer groups (those groups that were presented with totally different letter sets during learning and testing) or 'Scrambled' Transfer groups (the two groups that were presented with the same letter set during the two phases of the experiment, but the positions of the letters within the grammar were 'scrambled' for presentation during testing), as well as no significant differences between the two Control conditions or the different testing phase symbol sets. However, when the data were collapsed across both Control conditions and all four Transfer conditions, Transfer subjects (Mean=56.3% correct) were shown to be significantly less accurate in making well-formedness decisions as compared to Control subjects (Mean=64.6% correct), $F(1,84)=47.7$, $MSe=27$, $p < .01$. Although this difference exists between Transfer and Control groups, subjects in these respective conditions were still performing significantly above chance, as determined by a Classical Normal Approximation to Binomial test (Transfer subjects: $z=9.7$, $p < .001$; Control subjects: $z=16.0$, $p < .001$). Therefore, although Transfer subjects were performing significantly poorer than Controls, they still picked up the structure of the grammar and were able to transfer it across different stimulus forms. Any speculation as to the theoretical significance of this finding cannot be made with these data, however. This stems from the fact that the percent correct scores generated from subjects' Yes/No decisions are very likely influenced by a response bias, and this possibility must be eliminated before accuracy on the well-formedness task can be used to theorize on how implicit knowledge is represented. The controlling of response bias is accomplished by measuring subjects' sensitivity in responding to test items; this is accomplished by obtaining d' scores, and the analysis of these scores is discussed later.

Reaction Times

Theoretical Considerations. One variable which can possibly provide more insight into the issue of the representation of nonconscious information is the speed with which well-formedness judgments are made. If either abstract or distributive components are dominant in a tacit system, certain patterns of reaction times (RT's) should emerge. Theoretically, if knowledge was represented in an exemplar-based or fragmentary form, Transfer subjects would have to convert presented test items to match training items in some type of analogical process, resulting in Transfer subjects' RT's being longer than Control subjects' RT's. As mentioned earlier, such a process has recently been suggested by Brooks & Vokey (1991), who argue that nonconscious transfer across orthographic forms occurs via subjects forming 'abstract analogies'. To provide an example, if one studied the item TXXS and was tested on the item LBBW, transfer would occur by the subject reasoning that 'L replaces T, B replaces X, and W replaces S'. On the other hand, if tacit knowledge is primarily represented by abstract prototypes of rules, such a conversion would not be necessary, since the initial stages of the representation's formation would result in the formation of such a prototype. For example, if the item TXXS was initially studied, it would be converted into the prototype 'One occurrence of symbol 1, followed by 2 occurrences of symbol 2, ending with 1 occurrence of symbol 3'. If the item LBBW was later encountered, it would map directly onto the existing prototype, with no conversion process necessary. If subjects represent tacit knowledge in the form of abstract prototypes, there should be little, if any, differences between Transfer and Control groups' reaction times.

Initial Comparisons: Control, 'Regular', & 'Scrambled' Transfer. Initially, the same comparisons made for percent correct of 'Regular' and 'Scrambled' Transfer groups, as well as the two Control groups, were carried out. 'Regular' Transfer groups did not differ from one another; this finding was replicated in comparing the two Control groups. However, the A-A* group was found to be making grammaticality decisions slower than the B-B* group. A closer inspection of the raw data revealed this difference to have been due to 5 subjects in the A-A* condition that were responding with RT's twice as long as the remaining 10 subjects in the group, for no explainable reason. As these 5 subjects were not different in any other way from the remaining subjects in the group, these outliers were removed from the analysis, which eliminated the difference between the two 'Scrambled' Transfer groups, as well as the difference between the 'Regular' and 'Scrambled' Transfer groups. This equality in RT's between the two types of Transfer groups suggests that abstract elements may play a more dominant role in the synergistic relationship between these elements and their more exemplar-based counterparts. A distributive account of these data would suggest that RT's for 'Scrambled ' Transfer groups should be slower than RT's for 'Regular' Transfer subjects, as the 'Scrambled' items must be first be "undone" in order to understand their relationship to training items. Such a process seems intuitively more difficult (and longer) than plainly mapping one surface form to another, yet the data do not reflect this explanation.

Main Analyses. To explore any additional implications of RT's on implicit memorial systems, several ANOVA's were carried out on the RT data. With the five outliers removed from condition A-A*, Transfer and Control groups were collapsed across each other. The first test looked at Symbol Manipulation (Transfer or Control) x String Type (G and NG) x Block (1 and 2), while the second looked at Symbol Manipulation x Judgment Accuracy (Correct and Incorrect) x Block. The results indicated several interesting findings. Initially, there was a main effect of Manipulation, as Transfer subjects took significantly longer to make their well-formedness decisions, $F(1,81)=6.9$, $MSe=13.5$, $p < .05$, with a mean decision latency of 5.933 seconds, compared to Control latencies of 4.836 seconds. This seemingly lends support to the distributive/fragmentary view on the issue of nonconscious transfer. However, breaking down the significant Manipulation X Block interaction, seen in Figure 3, provides additional information that reveals that this main effect may not mean what it appears to on the surface. As can be seen in the Figure, although both Transfer and Control subjects exhibited faster RT's during the second block of testing trials (when compared to block 1), Transfer subjects showed a much greater rate of improvement ($F(1,81)=8.8$, $MSe=1.4$, $p < .01$). On the average, Transfer subjects made their well-formedness judgments 1.199 sec.

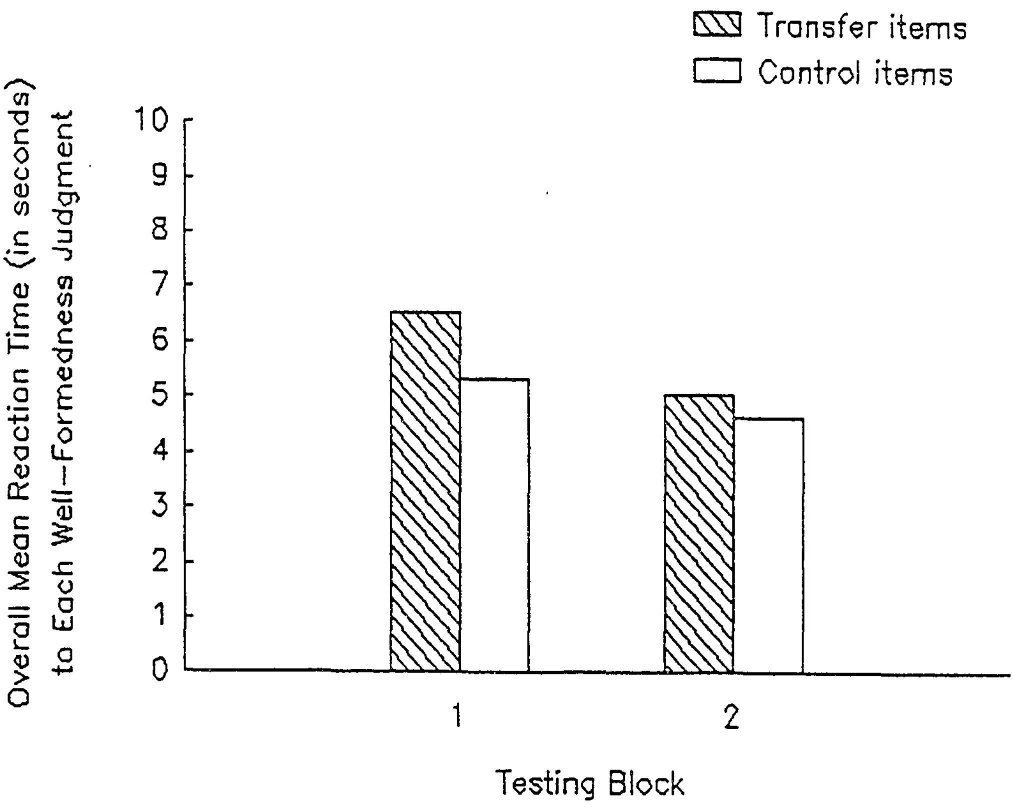


Figure 3. Mean reaction time (in seconds) to well-formedness judgments for Transfer and Control items during Experiment 1, as a function of test block.

faster on the second block of trials, while Control subjects improved their latencies by only .401 seconds. If this interaction were due to practice, then there should have been a significant block effect for the percent correct data. Since this did not occur, it seems that the difference between the Transfer and Control groups in terms of the overall RT may be explained by the manner of the design of the study. With subjects being tested solely on Transfer or Control items, as opposed to making well-formedness decisions for both types of stimuli, they may engage in processing that is initially limited to exemplar-based components. Only after a certain amount of time does processing become more abstract, an assumption that gains some support from the significant Manipulation X Block interaction.

As with the better accuracy of subjects in judging Grammatical items being due to similarity between G items and training stimuli, the similarity of G items to training stimuli had an effect on decision latency, as Grammatical items (Mean RT=5.205 sec.) were responded to at faster latencies than NG items (Mean RT=5.564 sec.), $F(1,81)=16$, $MSe=.6$, $p < .001$. Since subjects are not exposed to NG items during learning, they are less familiar with these items. As a result, subjects take longer to respond to these items during testing because Nongrammatical items are not necessarily part of one's exemplar-based representation in the early stages of the formation of a nonconscious representation. It may be the case that implicit representations are initially distributive and/or fragmentary in nature, becoming abstract only after extended practice.

In addition to these findings, RT's during the second 50 testing trials were significantly faster than RT's during the earlier portion of the testing phase ($F(1,81)=36.7$, $MSe=1.4$, $p < .01$), with RT's being 5.787 vs. 4.983 secs., respectively. The ANOVA found several significant results, displayed in Table 1. First, although

Table 1

Mean reaction times (in seconds) for correct and incorrect well-formedness judgments during Experiment 1, across Transfer and Control items and two testing blocks

Judgment Accuracy	Letter Set Manipulation		Block		Means
	Transfer	Control	1	2	
Correct	5.954	4.718	5.666	5.006	5.336
Incorrect	5.907	5.084	6.026	4.965	5.496

subjects responded equally fast when making correct and incorrect well-formedness judgments (F=1.1, p > .25), there were significant interactions of Manipulation x Accuracy (F(1,81)=6.3, MSe=.5, p < .02) and Block x Accuracy (F(1,81)=6.1, MSe=.5, p < .02). As can be seen in Table 1, the first interaction seems to be due to the fact that Control subjects had slower latencies for incorrect judgments, while Transfer subjects exhibited virtually no difference in RT's for correct and incorrect decisions. As for the Accuracy x Block interaction, this was apparently due to the fact that although RT's decreased across both blocks for both correct and incorrect decisions, incorrect decisions were made 1061 msec faster during block 2, while correct judgments were reached 660 msec faster during the second 50 testing trials.

Taken together, the analyses of reaction times extend the suggestions made earlier that nonconscious knowledge is represented in both exemplar-based and abstract forms. It appears to be the case that during the early stages of the formation of a representation, distributive components guide the representational process, while abstract elements play a stronger role as one's memorial system becomes more complex.

Confidence Ratings

Since subjects were never informed as to the accuracy of their well-formedness judgments, confidence ratings provide an insight into subjects implicit awareness of their well-formedness performance. To see if any differences existed between the groups, a one way ANOVA was performed and found a significant main effect of group, $F(5,84)=3.4$, $MSe=1.9$, $p < .01$. As can be seen in Table 2, this difference is clearly due to the lower confidence ratings of the A-A* group. There appears to be no clear reason for this outcome, but if one looks at the remaining Transfer groups against the Control groups, the difference are minuscule. It seems that no matter how stimuli appear on the surface, subjects are equally confident in making well-formedness decisions.

To see if there were any other confidence effects, the same analyses conducted on the RT data were carried out on the confidence data, and the main results are also displayed in Table 2. Subjects were found to rate their confidence to Grammatical items higher than

Nongrammatical items ($F(1,84)=49.3$, MSe=.2, $p < .001$) and correct decisions higher than incorrect decisions ($F(1,84)=18.4$, MSe=.1, $p < .01$). Since subjects were never informed as to the accuracy of their well-formedness decisions, this latter finding suggests that subjects know something about their decisions, although they may not be able to verbalize much of that knowledge. The only other significant

Table 2

Mean confidence ratings for each group during Experiment 1, across correct and incorrect judgments, & Grammatical and Nongrammatical items

	Judgment Accuracy		String Type		
Study-Test Symbol Sets	Correct	Incorrect	G	NG	Means
A-A	5.43	5.09	5.44	5.08	5.26
B-B	5.36	5.15	5.49	5.03	5.26
A-B	5.36	5.26	5.46	5.16	5.31
B-A	5.28	5.15	5.35	5.09	5.22
A-A*	4.54	4.41	4.62	4.34	4.48
B-B*	5.32	5.23	5.50	5.06	5.28
Means	5.21	5.06	5.31	4.96	5.14

finding was a Manipulation x Accuracy interaction ($F(1, 86)=5.2$, $MSe=.1$, $p < .03$), but as with the group effect, this interaction is due to the inclusion of the lower A-A* confidence ratings in the mean Transfer scores.

Sensitivity Analysis

To determine the transfer manipulation's effect on subjects' sensitivity to making their grammaticality decisions, several signal

Table 3

Mean d' and c scores for each group during the well-formedness phase of Experiment 1

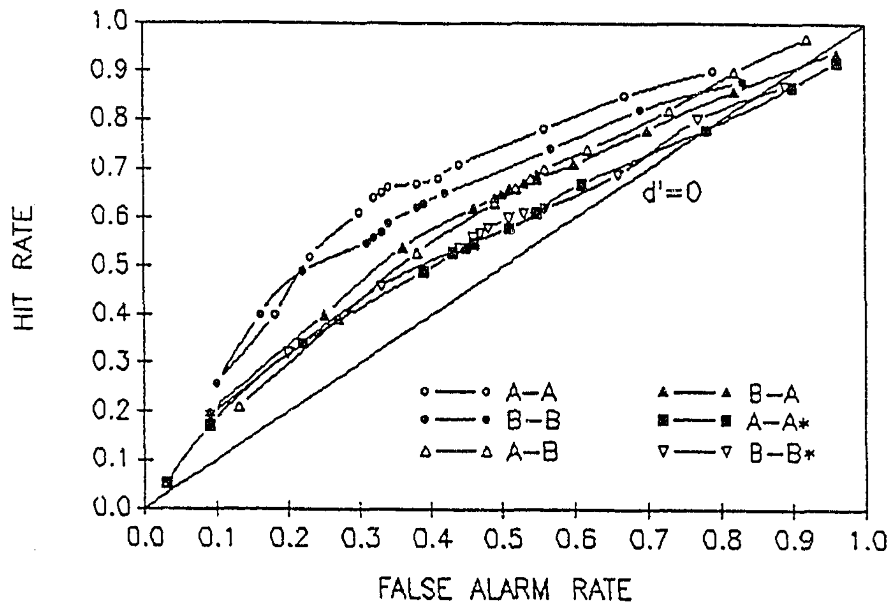
Study-Test Symbol Sets	Statistic	
	d'	c
A-A	.71	.03
B-B	.58	-.07
A-B	.36	.24
B-A	.32	.22
A-A*	.16	.02
B-B*	.22	.09

detection analyses were carried out. Initially, d' scores from each group were compared, and as can be seen in Table 3, there was a significant Group effect, $F(5,77)=29.9$, $MSe=.06$, $p < .0001$. Newman-Keuls tests found this effect being mainly due to higher d' scores being found for Control subjects as compared to the means of the four Transfer conditions, as well as significant differences between the two 'Regular' and two 'Scrambled' Transfer conditions, all p's < .05.

The better sensitivity of Control subjects as compared to Transfer subjects does allow speculation into how nonconscious knowledge is represented. Unfortunately, each of the three implicit representation perspectives (abstract, exemplar-based, and fragmentary) can theoretically explain the obtained results. However, one analysis which could make a differentiation between the three perspectives is accomplished by looking at the results from the two 'Scrambled' Transfer conditions. If one's representation was exemplar-based, then one would rely heavily on the individual letters and their positions within the grammar to make grammaticality decisions. However, an abstract representation would presumably be more sensitive to the underlying structure of stimuli, independent of the letters and their positions. With these alternatives in mind, if the positions the letters assumed in the grammar were changed, as they were in the 'Scrambled' Transfer conditions, an exemplar-based system would have difficulty in classifying Transfer items; performance here would likely be at chance since there is no representation to accomodate the new items. An abstract representation, however, would allow for classifications to occur above chance, as one's representation of the grammar's structure could be applied to the 'Scrambled' items. In accordance with this latter assumption, the obtained d' scores from the two 'Scrambled' Transfer groups reveal judgment sensitivities that were above chance, indicating that while exemplar-based knowledge may play a role during the initial formation of as tacit system, abstract elements come into play later on to account for transfer.

In addition to the d' scores, ROC curves were generated for each group by plotting the Hit and False Alarm rates of 'Yes' and 'No' responses made to G and NG items at each (1-7) confidence level. Z-scores of the Hit and False Alarm rates were also plotted. These curves, displayed in Figure 4 below, mimic the d' results, in that Control groups were more sensitive to stimuli, as compared to Transfer subjects. One interesting point from these curves is that although each group's responses appear to be normally distributed, both 'Scrambled' Transfer groups drop below chance responding as the Hit and False Alarm rates approach 1.

Finally, subjects' bias towards responding 'Yes' or 'No' was measured by obtaining a c statistic for each group, the results of which are also in Table 3. Interestingly, only the 'Regular' Transfer groups (A-B and B-A) exhibited a slight bias towards responding 'No'. This result is slightly puzzling, since it seems more intuitive that 'Scrambled' Transfer subjects would show a "No" bias as a result of their odd-looking stimuli.



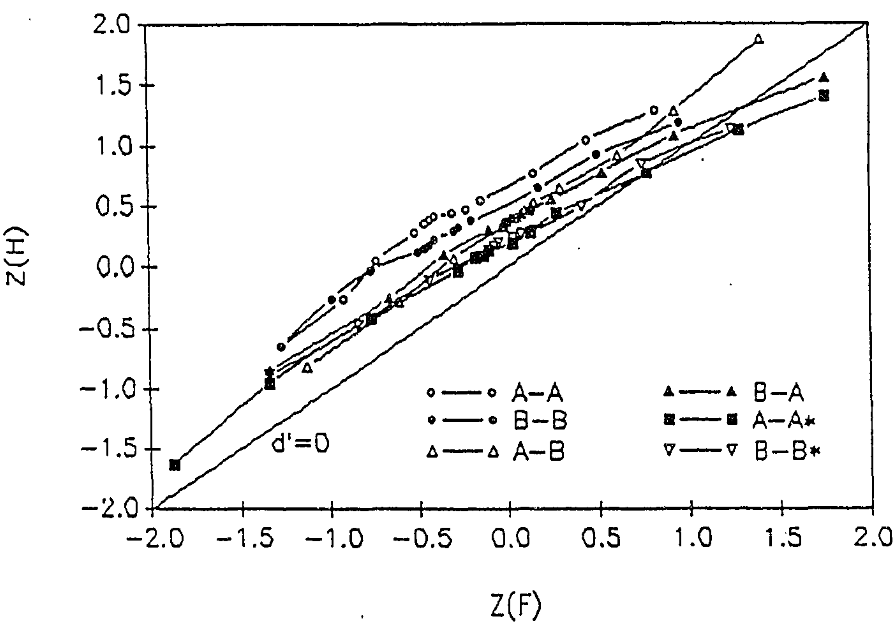


Figure 4. ROC curves for Hit and False Alarm rates (top), as well as z scores (bottom), for Experiment 1.

<u>Consistency Scores</u>

To measure the degree to which subjects form rules which are representative of the artificial grammar (Reber, 1989), consistency scores were measured by recording whether or not subjects responded correctly (C) or erroneously (E) to each letter string on each presentation of each item. Since each string was presented twice, CC scores reflect correct responding on both presentations, EC and CE scores, if significantly different from one another, tap the degree to which one's memory is improving or faltering, respectively, and EE rates reflect the amount of non-representative rules formed and utilized by subjects. If the EE rate is significantly higher than the average of the CE and EC rates, then one can infer that subjects were utilizing rules which did not correspond to the actual structure of the grammar.

To investigate for any possible consistency effects, a one-way ANOVA was conducted on the overall consistency score (arrived at by adding the CC and EE rates together). The results of this analysis, shown in Table 4, revealed no significant differences existing between the groups, $F < 1$, $p > .25$. Similarly, a planned comparison between the mean of the two Control groups and the mean of the four Transfer groups yielded no significant difference, $F < 1$, $p > .25$. These findings fail to lend conclusive support to any of the three positions on the form in which tacit knowledge is represented, as any of the three approaches could handle subjects using rules consistently. What is interesting, however, is that although there were no overall

Table 4

Consistency scores for each group during Experiment 1

Study-Test Symbol Sets	Consistency Score CC	CE	EC	EE	Total
A-A	54	13	12	21	75
B-B	48	15	16	21	69
A-B	41	14	16	29	70
B-A	44	13	15	28	72
A-A*	42	14	14	30	72
B-B*	41	14	14	31	72
Means	45	14	14	27	72

differences between the groups, there was evidence of inaccurate rule formation, as the EE rate significantly exceeded the mean of the EC and CE rates in every group, as measures by chi-square tests (all z's > 2.8, all p's < .01, two-tailed). In addition, while Control conditions resulted in EE rates that were lower than the mean EE rate of the four Transfer conditions (21 vs. 29.5, respectively), this difference was not significant, z=.76, p=.44 (two-tailed). Despite this result, the high EE rates suggest that subjects are 'inventing' non-representative rules to handle the cases where they do not know the status of an item, and they tend to use these non-representative rules consistently. What is likely occurring is that during the formation of an implicit knowledge representation system is that some knowledge is represented abstractly, some is exemplar-based, and still some is represented in a fragmentary form. As a result of such an eclectic distribution of knowledge forms, some information about the structure of the grammar may be inaccurate. For example, if the item TSXS is represented in its raw form, and the transfer item GTRT is encountered, one might be forced to invent a rule to deal with the grammatical status of the latter item because no knowledge about the raw form of this item is in the representation. While this explanation begins to get at the heart of implicitly acquired storage systems, less speculative information about the composition of an implicit representation was obtained by analyzing subjects' responses at a more microscopic level.

Microanalysis of Responses

An additional analysis which could provide insight into how subjects represent implicitly acquired information centers on subjects' overall accuracy with individual items. Such an investigation allows one to determine if subjects have a deep and complex representation of the stimulus environment, or if they represent tacit knowledge in a very limited format. If subjects represent the grammar by way of nothing more than a few specific instantiations of highly salient stimuli, then those items should be responded to with high levels of accuracy, and when those items are removed from the overall distribution, performance should drop off. However, if one's representation contained a rich assortment of complex information in addition to highly salient patterns, one would expect performance levels to remain fairly stable when salient items are removed from the distribution.

Saliency. Investigating item saliency, however, is a tricky matter, for what may appear to be salient to one subject may not be very salient to another. Therefore, before any statistical analyses were carried out, it was necessary to determine how saliency should be defined, a process which settled upon looking at the top and bottom 10% of the items for each group, in terms of overall percent correct. Those items are displayed in Table 5. As can be seen in the Table, the only recurring pattern of saliency across the groups seems to be that subjects respond very accurately to items that either have three or more of the same letters in a row (e.g., BMQQQQ), or items that have more than 1 repetition (e.g., MHXXMMQQ). Therefore, an initial analysis was carried out by separating these 'salient' items (n=20) from the remaining 'non-salient' stimuli (n=30). One-way ANOVA's were performed on salient vs. non-salient items in each condition, and in all 6 cases, as seen in Figure 5, there were non significant differences between the mean percent correct of salient vs. non-salient stimuli, all $F's(1,49) < 2.71$, all $p's > .1$. Therefore, although subjects might be able to notice and discriminate some items better than others, their representations contain much more than a few

Table 5

Items with individual percentages correct in the top and bottom 10% of all test stimuli for Experiment 1

Study–Test Symbol Sets	Proportion of Distribution: Top 10% (Pct.)	Bottom 10% (Pct.)
A–A	QHMXQQH* (97)	MXQ* (33)
	MHXXQBH (93)	MHHXH (33)
	MHXXMMQQ (90)	BMMMMMQQ (23)
	BMQQQQ* (90)	MHHHHXH (17)
	MXXMMQQ (90)	MXQBH* (13)
B–B	LZZLLRWJ (93)	LJJZJ (40)
	RJLZRRJ* (93)	WLRWZRWJ (37)
	LJZZLRWJ (90)	LZRWJ* (27)
	RWZLRR* (90)	WLLLLLRR (17)
	WRZWRZWZ* (83)	LJJJZJ (3)
A–B	LJJZZRWJ (93)	WZWRZRLL* (23)
	LJZZRWJ (87)	WLLWJ* (23)
	WRWZRR (87)	LJJJJZJ (20)
	LZZLLRWJ (87)	LZR* (20)
	LZJ (83)	WLLLLLRR (13)
B–A	BMMMQQ (90)	MXQ* (33)
	MXXMQQ (87)	MBMXH (33)
	MXXMMQQ (83)	MHHHHXH (30)
	MHXXMMQQ (80)	BMMMMMQQ (27)
	BMMQBH (77)	MXXQX* (23)
A–A*	MQXXXX* (93)	QMQHB* (13)
	QHHQQXMB (83)	MQQQQQXX (13)
	MXMHXX (83)	BHHXMB* (3)
	QBHHQQXX (80)	QBBBBHB (3)
	MXMHQXMB (80)	MQQQQXMB (3)
B–B*	RJJRRZLW (93)	RWWWWJW (27)
	RWJJRRZZ (93)	LRRRRZLW (27)
	LZJLZJLJ* (87)	WJJZLW (23)
	RRZZ* (87)	LRRRRRZZ (10)
	RJW (87)	RJZ (10)

Note. *'s indicate a Nongrammatical item.

salient patterns.

Item Length. One final microanalysis dealt with item length. It might be the case that subjects' representations are based upon items that are easy to identify because of their length. For example, the grammaticality status of the item LZJ might be easier to decide than the item LJZZRWJ because the former stimulus is shorter than the latter item. To investigate the possibility of such an effect, for each condition, the percent correct of stimuli of length 3-8 were compared with an ANOVA. The results, in Figure 5, supported the current hypothesis that tacit knowledge is represented by a complex data base, as string length failed to show a significant main effect in every condition, all F's$(5,49) < 1.33$, all p's $> .25$.

Interview

The interview session at the conclusion of the well-formedness task suggest what has been hypothesized elsewhere about the artificial grammar learning procedure (Reber & Allen, 1978): subjects, for the most part, fail miserably in reporting rules that reflect the actual structure of the grammar, yet they still perform significantly above chance in making grammaticality decisions. In general, the 'rules' that subjects reported as guiding their grammaticality judgments were very vague and intuitive, and did not contain much specific information about the structure of the grammar (e.g., "I looked for patterns", or "The letters just looked right or wrong"). When subjects did report specific 'rules', these descriptions were either

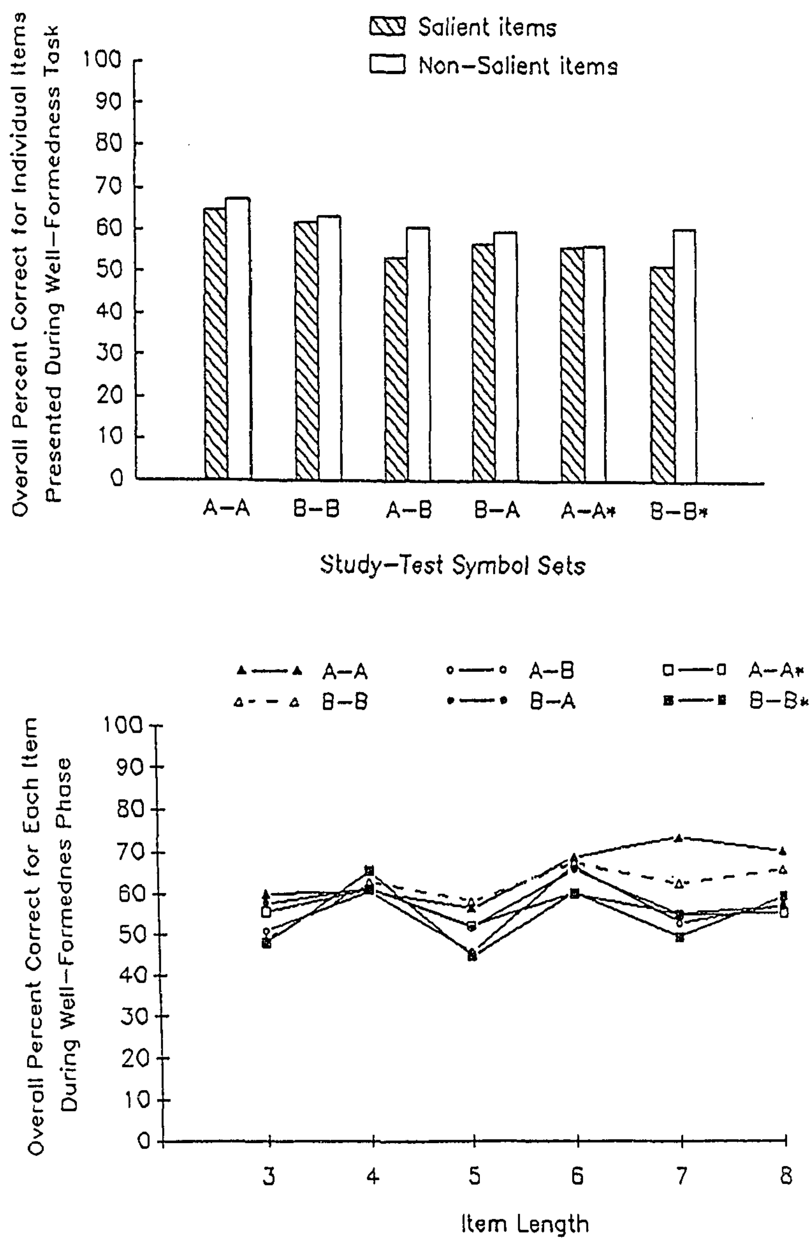


Figure 5. Percentages correct of salient and non-salient items (top) & items of each string length (bottom), from Experiment 1.

inaccurate, e.g., "J's couldn't appear more than twice in a row" or specific to the beginning or terminal states of the grammar, e.g., "M could appear in the first position". Very rarely did subjects report conscious knowledge of the inner states of the grammar.

It seems to be the case that subjects, across all conditions, seem to form what Dulany, et al. (1984, 1985) called 'correlated' or 'local' grammars. Such grammars reflect fragmentary aspects of the rule structure of the grammar used to generate stimuli and enable subjects to function fairly efficiently in the stimulus environment known as the artificial grammar learning procedure. These local grammars can take on several forms, with the most common being, for example, statements such as 'If the item began with the letters F and S, it's grammatical'. Such grammars are perfectly fine for the current argument that tacit representations contain both specific and abstract information, for although one might know the specific elements of the beginning of an item, they may have an abstract understanding of the remainder of that item.

As for the Transfer subjects specifically, the reports that were of most interest were if and how they were able to understand the relationship between the different learning and testing letter sets. Subjects primarily reported general rules (e.g., "I looked for patterns/comparisons"), and when asked about drawing analogies between the two letter sets, they either felt, overall, that "the task was very confusing" or were only able to report analogies for the initial and terminal states, e.g., "M's replaced L's".

Summary

Overall, the results from this experiment are ambiguous at best. Although the Transfer groups did perform slightly worse in terms of their overall d' and percent correct scores, their performance was still significantly above chance, but it is not clear whether this performance was due to the formation of abstract analogies or a prototype of the rules of the grammar, or a combination of the two. The RT data, which could provide additional insights into the representation issue, were not reliable enough to formulate concrete conclusions. The next study was carried out in an attempt to clarify these issues.

Experiment 2

The positive transfer results from the first experiment did not provide sole support to the abstract, distributive, or fragmentary positions. As a result, several issues remain concerning the nature of implicitly acquired representations. First, although subjects who were presented with stimuli from different or rearranged letter sets performed above chance on the well-formedness task, their performance was significantly poorer than that of Control subjects. In addition, the results of the first experiment failed to adequately determine if subjects form an abstract prototype of the stimulus display, facilitating transfer, or use some type of analogical reasoning process to decide upon the grammatical status of Transfer items, as proposed by Brooks & Vokey (1991). Although the Transfer groups exhibited significantly slower RT's as compared to Controls (a finding which would support the distributive explanation of the transfer process, the existence of knowledge of 'abstract analogies' was not specifically tested in Experiment 1.

Experiment 2 addresses these issues as follows. The poorer performance of the Transfer subjects in relation to the Control subjects in the previous study may have been due to a processing deficit of some sort. Either the learning phase was not as attention-demanding as necessary, with the result being that the subjects may have failed to detect some of the positional covariances, or subjects failed to engage in transfer-appropriate processing (TAP). This latter process is defined as cognitive operations that are engaged in during learning that facilitate processing on a secondary, transfer, task (Graf & Ryan, 1991). Manza (1992) has suggested that TAP during the learning portion of an artificial grammar learning task facilitates performance on a subsequent testing phase, while the lack of TAP interferes with the implicit transfer process. To investigate these issues, the learning phase of the next experiment was altered to increase the attentional demands required by subjects in order to successfully complete the learning task. To provide additional insights into the representation issue, Experiment 2 employed a within-subjects design. Instead of exposing subjects to just one letter set during testing, all subjects were tested on both a Control and Transfer letter set during the well-formedness phase. In addition, at the conclusion of the well-formedness task, subjects completed a matching task, the results of which should lend better insight into the nature of a tacit knowledge representation. Finally, the testing phase consisted of a total of 400 trials: 200 Control strings and 200 Transfer strings. This feature of the design, when coupled with the more demanding learning phase, should eliminate the RT differences found in Experiment 1 and allow for a clearer analysis of the content of a tacit knowledge representation.

Method

Subjects, Design and Materials

Twenty Brooklyn College students were solicited from various introductory and upper-level psychology courses to participate in this study. Subjects were compensated at the rate $5.00 per hour or received course credit for their participation. In addition, a performance incentive was implemented, whereby subjects would receive $.05 for each well-formedness judgment they made that was above chance performance (50%) on the well-formedness task.

The lengthy within-subjects design implemented here necessitated the use of a grammar, shown in Figure 6, that is radically different from the one used in Experiment 1. The reason for utilizing this new grammar (from Brooks & Vokey, 1991), is that the number of stimuli needed for presentation was more than the grammar from Experiment 1 could generate between 3 and 8 letters in length.

Learning. For the learning phase, 50 items from letter set A were selected for presentation. The reason for not presenting learning items from letter set B to control for any letter set effect was that (1) Brooks & Vokey (1991) found no letter set effect using this grammar,(2) there was no effect of letter set in Experiment 1, (3) Mathews, et al. (1989) used various letter sets and found no letter set effect independent of a transfer manipulation, (4) Manza (1992) controlled for three different letter sets and also found no letter set effect. It would seem redundant to control for this again since it has become obvious that the letters used to instantiate an artificial grammar, independent of any other manipulation, do not effect performance.

Testing. Fifty new Grammatical items were selected for presentation during testing, plus another 50 Nongrammatical items.

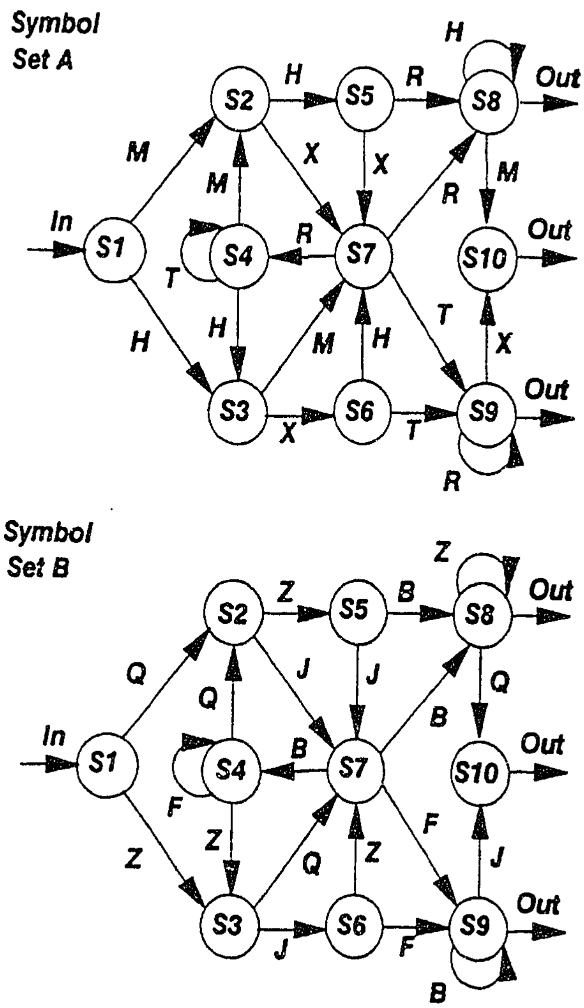


Figure 6. The artificial grammar used to generate stimuli for Experiment 2, in its two instantiations.

The test items generated using letter set A were also instantiated in letter set B, so each item was presented in both its letter set A and B forms (see Appendix C for a listing of the specific stimuli). To control for order effects, subjects were randomly assigned to one of two groups: (1) those who first received a block of Control stimuli (letter set A) during testing, followed by a block of Transfer stimuli (letter set B), and (2) those who initially received Transfer stimuli followed by Control stimuli.

Procedure

Learning. The acquisition phase differed from that in Experiment 1 by changing the learning criterion. In order to increase attention, and, in turn, facilitate some form of TAP, each presentation contained 2 letter strings from set A, with one item above the other. Each item in each learning pair was randomly selected, without replacement, from the pool of 50 items. Therefore, subjects were initially exposed to 25 pairs of items, with each pair presented for seven seconds. The criterion for each pair was to correctly reproduce both letter strings on the same exposure. Once the criterion was met for all twenty-five pairs, the 50 items were randomized in different pairings, and subjects proceeded through a second block of presentation/reproduction trials. Once this was completed, they were given a brief (2-3 minute) rest interval, after which the instructions for the next phase were administered.

Testing. Following instruction about the rules of the grammar and the nature of the Transfer stimuli, subjects were presented with a total of 400 trials, in the following order: subjects in group AB were presented with 50 G and 50 NG Control items, followed by 50 G and 50

NG Transfer items, followed by the same 100 Control strings, and finally the same 100 Transfer strings. As for group BA, the same general alternation of Transfer and Control items occurred, but the alternating items began with Transfer stimuli. In addition, although Transfer and Control stimuli were matched, the items in the Transfer set were presented in a different order than Control items. Following the post-experiment interview, subjects were administered a matching task, where they had to match up the Control letters with their Transfer counterparts.

Results and Discussion

Learning Phase

Trials to Criterion. The data of interest from the acquisition phase were the mean number of trials to criterion for each study pair. As all subjects received the same learning procedure, the data were subjected to an ANOVA investigating performance across the first and second block of 25 learning pairs. Since presentation order was randomized, the hypothesis that subjects actually learn the rules of the grammar during the study phase could be investigated. The results revealed that subjects performance improved as testing progressed. Subjects required a mean of 6.9 presentations per pair to reach criterion during the first block of 25 pairs, compared to a mean of 5.0 presentations per pair during the second block, an improvement which was significant, $F(1,18)=26.8$, $MSe=6.8$, $p < .01$.

Testing Phase

Percent Correct. The different ordering of Transfer and Control items did not have any effect on judgment performance, as the Group effect failed to reach significance at the .05 level. Although it was originally hypothesized that making the learning phase more demanding (attention-wise) might increase the performance with Transfer items as a result of subjects forming a more abstract representation, this did not occur, as a Group X Symbol Manipulation X String Type X Block ANOVA found a significant difference between subjects' overall percent correct for Transfer (54.3%) and Control (60.8%) strings. Although both of these measures significantly exceeded chance levels at the .01 level (Transfer z=5.4, Control z=13.6), the implementation of a more complex grammar led to results that replicated those from Experiment 1, in that the percent correct data fail to conclusively support any of the three perspectives on the representation of tacit knowledge. However, the data do seem to suggest that increasing the load on memory may not necessarily alter the nature of a tacit representation.

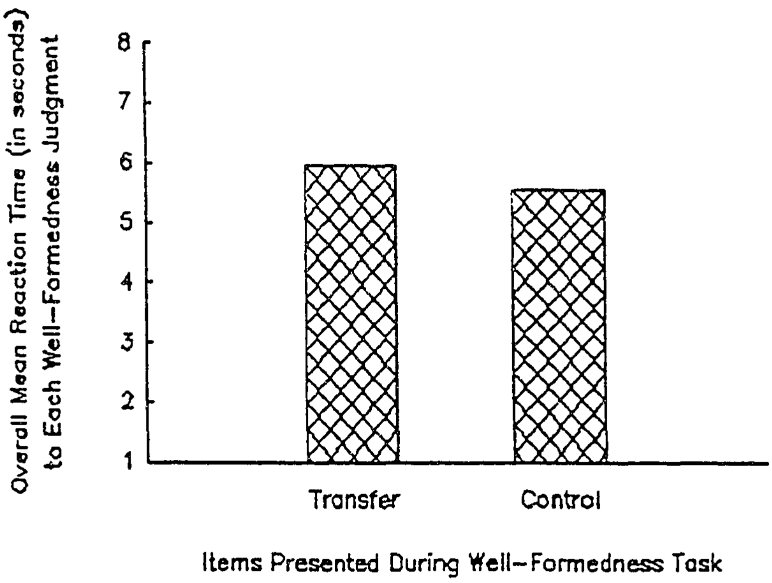
Matching Task. Major proponents of the distributive view (Brooks & Vokey, 1991) have argued that implicit transfer performance can be explained via subjects forming 'abstract analogies' of the stimulus environment. These relationships basically involve the replacing of letters from one letter set with letters from a new letter set. For example, if the transfer item LZZ is encountered during testing, and one has a stored representation of MQQ, a grammaticality judgment for the former item can be made by reasoning that "L replaces M, and Z replaces Q". Although Brooks & Vokey (1991) never tested this notion empirically, if this kind of abstract analogy is indeed being formed, subjects should be able to identify which letters from the transfer set are paired up with specific letters from the training set. To test this notion, a matching task was administered to all subjects immediately after the conclusion of the testing phase. The task for subjects was to pair individual letters from the training set with their transfer set counterpart. A one way ANOVA found no significant difference between the AB and BA groups, so the data were collapsed across the two groups and subjected to a Classical Normal Approximation to Binomial Test. With five letters in each set, chance performance for this task was .2, and the overall mean performance of 1.4 correct matches (out of a possible 5) was found to not differ significantly from guessing, $z=.187$, $p > .2$. Although this test reflects conscious knowledge of the relationships between transfer and control symbols, and the present discussion is focused on nonconscious knowledge, this finding still suggests that 'abstract analogies' may not have as strong an effect on classification performance as Brooks & Vokey might suggest. Data from subjects reaction times to their grammaticality decisions, discussed below, slightly clarify this issue, providing additional information in the current debate over which process (abstractive or distributive) may play a more primary role in the transfer and representation of nonconscious knowledge.

Reaction Times. With the idea that RT's can provide insight into how implicitly acquired knowledge is accessed and represented, two 4-way ANOVA's were conducted on the RT data. As in the first experiment, the analyses looked at a) Group X Symbol Manipulation X String Type X Block and b) Group X Symbol Manipulation X Accuracy X Block. Initially, the presentation order of Transfer and Control items was found to have no significant effect on response latency, $F(1,18)=1.7$, $p > .2$. However, there was a significant main effect of Block ($F(1,18)=13.3$, $MSe=7.5$, $p < .01$), as RT's during the first 200 trials averaged 6.552 seconds, compared to mean latencies of 4.967 seconds for the second 200 trials. These faster Block 2 latencies may be based upon any of several assumptions. First, the effect may be simply a practice effect, for as time progresses and subjects understand the stimuli better, their decision latencies may decrease. Alternatively, arguing for a more synergistic relationship for abstract and exemplar-based knowledge, it may be the case that as time progresses, nonconscious representations become highly prototypical, allowing judgments to occur much more rapidly than early decisions, or that subjects' knowledge base becomes more loaded with specific instances, allowing faster processing of similarity factors.

A result that can clear up this uncertainty is the RT data for Transfer and Control items, shown in Figure 7. If the representation developed during an implicit learning procedure is abstract, then latencies for Transfer and Control items should be about equal. As above, the reasoning here is that if one has a prototype of the grammar stored in memory, there would be no need to convert Transfer items to the Control letter set. If the representation is distributive, however, this conversion would be necessary, resulting in longer response latencies for Transfer items, compared to Control stimuli. Recall, that for the previous experiment Control strings led to faster RT's, and it was argued that subjects may have not been exposed to enough information about the grammar to form a complete representation. The testing phase was lengthened in the current study in an attempt to eliminate this difference. In agreement with the abstractive position, Transfer and Control strings were responded to at about the same speed, $F(1,18)=1.8$, $p > .18$.

There were only two remaining important RT results, the first of which is also in Figure 8 and centers on the fact that Grammatical and Nongrammatical items were responded to with approximately equal latencies, $F < 1$, $p > .7$. Finally, although subjects were never explicitly informed as to the accuracy of each well-formedness decision, they responded significantly slower when they made an incorrect judgment as compared to when they judged strings correctly (6.022 and 5.572 seconds for incorrect and correct judgments, respectively), $F(1,18)=13.3$, $MSe=.6$, $p < .01$.

Confidence ratings. As is Experiment 1, subjects' confidence ratings were analyzed in the hope of finding insights into subjects grammaticality decisions. The confidence data were subjected to the same analyses as performed on the RT data. As in Experiment 1, subjects seemed to intuitively 'know' the accuracy of their well-



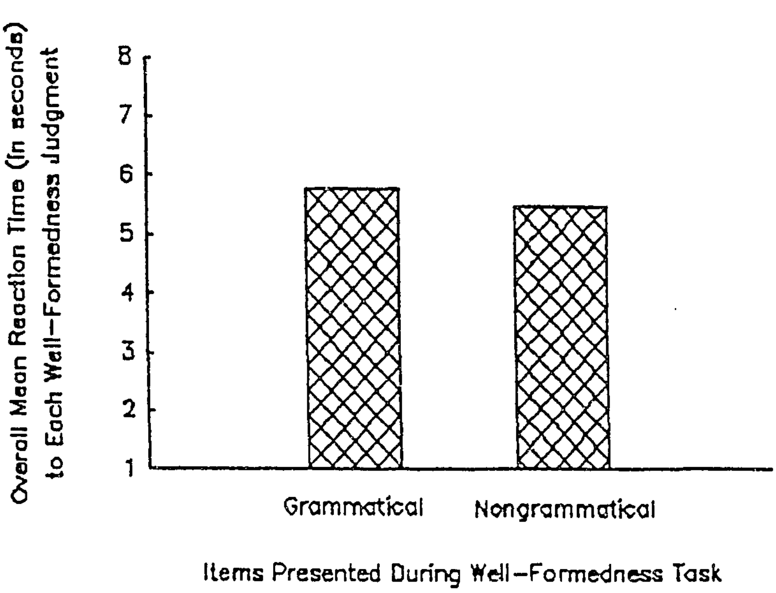


Figure 7. Mean reaction times (in seconds) to well-formedness judgments for Transfer and Control items (top), & Grammatical and Nongrammatical items (bottom) during Experiment 2.

formedness decision, as they felt significantly more confident following correct decisions as compared to their confidence levels after incorrect judgments (4.88 versus 4.67, respectively), $F(1,18)=5.1$, $MSe=.09$, $p < .01$. In addition, subjects reported approximately equal confidence ratings for G and NG items, as the String Type main effect failed to reach significance, $F(1,18)=2.2$, $p > .15$. Finally, the initial learning of Control items had some effect on confidence as compared to Transfer items, as Transfer items led to significantly lower confidence ratings than Control items (4.7 versus 4.85, respectively), $F(1,18)=5.9$, $MSe=.2$, $p < .05$.

Sensitivity Analysis. As in Experiment 1, subjects hit and false alarm rates were computed for each confidence level. As can be seen in Figure 8, although Control and Transfer ratings were normally distributed, subjects were more sensitive to Control items as compared to Transfer items. This difference was significant, as the d' scores (see Table 6) were significantly higher for Control items, $F(1,25)=115.6$, $MSe=.02$, $p < .001$. This result replicated the percent correct result, suggesting that an increased memory load has no significant bearing on decision sensitivity. Finally, there were no response bias' for either Transfer or Control items, as the c scores for both stimuli fell within the distribution of d' scores.

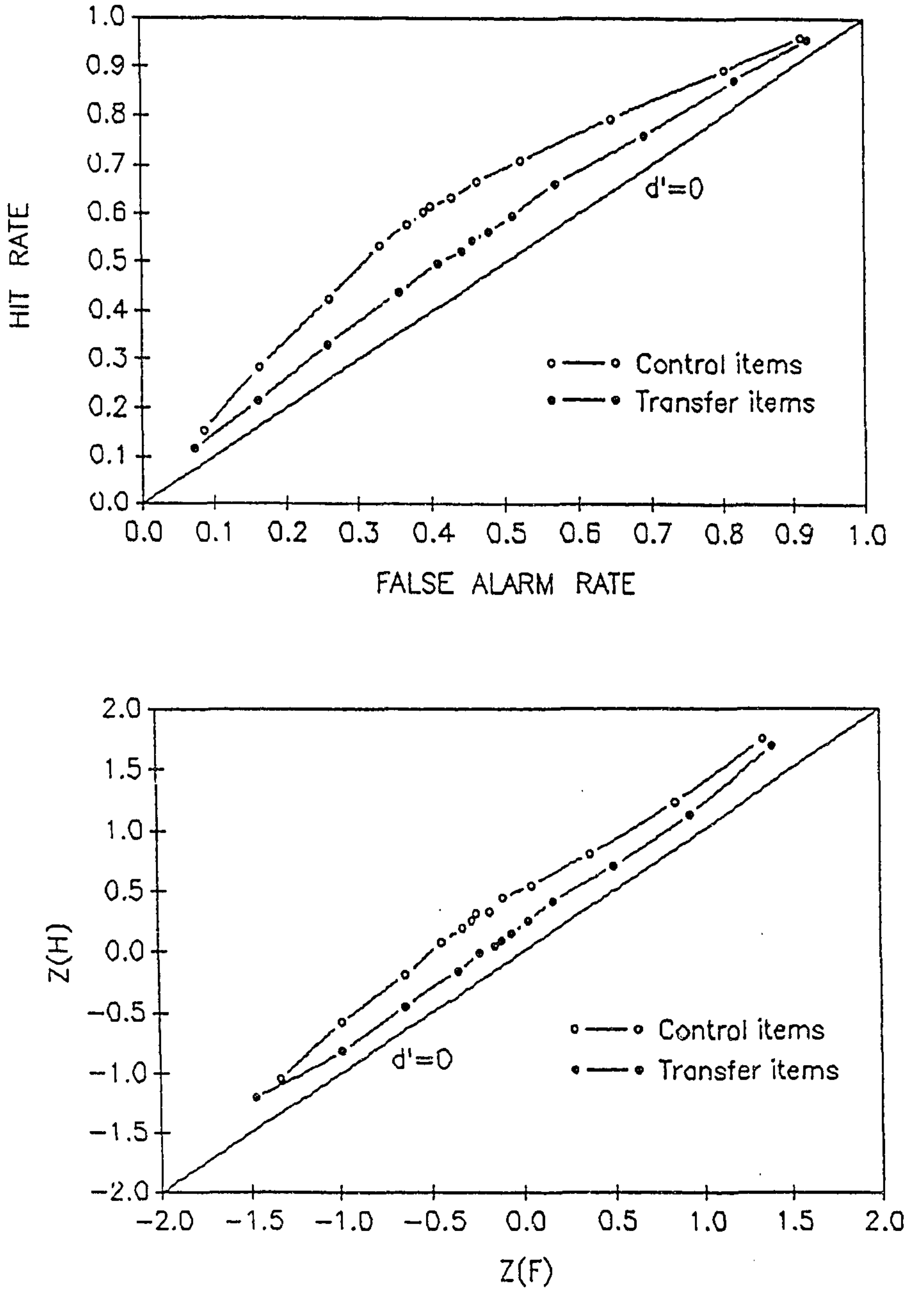


Figure 8. ROC and z curves for each group during Experiment 2.

Table 6

Mean d' and c scores for Transfer and Control items during the well-formedness phase of Experiment 2

	Statistic	
Test Stimuli	d'	c
Transfer	.22	0
Control	.47	.01

Consistency Scores. The consistency scores from this experiment, displayed in Table 7, offer several interesting findings. First, although there was no significant difference in the overall consistency scores of the two ordering conditions ($F < 1$), as well as between Transfer and Control stimuli ($F < 1$), there was a significant Stimuli x Test Order interaction, $F(1,18)=11.7$, $MSe=23.5$, $p=.003$. It seems that subjects who judged Control items first had higher overall consistency scores for Transfer items, while those subjects presented with Transfer items first were more consistent with Control stimuli.

There was also evidence that the knowledge subjects acquired was not a totally accurate representation of the stimulus environment, as the EE scores exceeded the mean of the CE and EC scores in every

Table 7

Consistency scores for each group during Experiment 2

Test Order-Stimuli	Consistency Score				
	CC	CE	EC	EE	Total
Control/Transfer					
Control	44	20	17	19	63
Transfer	39	16	15	30	69
Transfer/Control					
Control	44	17	15	24	68
Transfer	36	18	18	28	64
Means	41	18	16	25	66

cell except for Control items for those subjects studying Control items first (z=.79, p=.43, two-tailed; all remaining z's > 3.0, p's < .01).

Microanalysis of Responses. As in Experiment 1, the idea that tacit knowledge representations contain highly complex components, and not simple elements was explored by looking at the Percent Correct scores for individual items in several microanalyses. As can be seen in Table 8, the same pattern of salient items in the top and bottom 10% of the distribution found in Experiment 1 was found here, as subjects judged the well-formedness of items with repetition

Table 8

Items with individual percentages correct in the top and bottom 10% of all test stimuli for Experiment 2

	Proportion of Distribution	
Stimuli	Top 10% (Pct.)	Bottom 10% (Pct.)
Control	MXTT* (98)	MXRHMRHR* (40)
	HMTRMRRX* (85)	HMTRRX (40)
	MXRTTMXR (85)	MHXT (38)
	MMRH* (83)	MXRH (35)
	HXHRHMRH (83)	RHXRMHXR* (35)
	XXTRRX* (83)	MXRTHHHM* (35)
	HXHRMHXR* (80)	MHXRXMHR* (33)
	HHHRMXRH (80)	HXHTMXR* (30)
	MHRHHHHM (78)	HXHRHMXT* (30)
	HMRTHMR (78)	HXHRHXRH* (30)
Transfer	QJFF* (83)	QJBZ (35)
	ZJZZBBB (83)	ZJZBZQJF* (35)
	QQBZ* (83)	QJBZQBZB* (35)
	QJBFFZQQ* (80)	QJBQJFBB (33)
	ZQBFZQB (78)	BZJBQZJB* (33)
	ZZZBQJBZ* (78)	ZQBZQFJB* (30)
	ZZJBFQJB* (75)	QJBFZQF (28)
	QQBZZZQ* (75)	QZJBJQZB* (25)
	ZJZF (75)	ZBZF* (25)
	ZQBFFZQB (75)	ZJZBZJBZ* (23)

Note. *'s indicate a Nongrammatical item.

in several places with each item better than other types of items. In comparing these highly salient items to the rest of the distribution, shown in Figure 11, there was no significant effect of saliency ($p > .1$), as the overall percent correct of salient items was approximately equal to the accuracy with non-salient items. Lastly, items of length 3 through 8 were compared to determine if subjects could base their representation on shorter, and presumably, easier to process, stimuli. However, the results from Experiment 1 were replicated, as can be seen in Figure 9. There was no significant effect of string length, as subjects judged Transfer and Control items of all lengths with relatively equal accuracy, p's > .25.

Verbal Reports. The verbal report data from this study were basically the same as the reports from Experiment 1. Subjects once again failed to provide many specific, detailed, rules that accurately reflected the actual structure of the grammar, although they did provide evidence of their own local grammars. However, as evidenced by the high EE consistency scores described above, these personal grammars likely contain information that does not reflect the actual nature of the grammar used to generate the stimuli for this study.

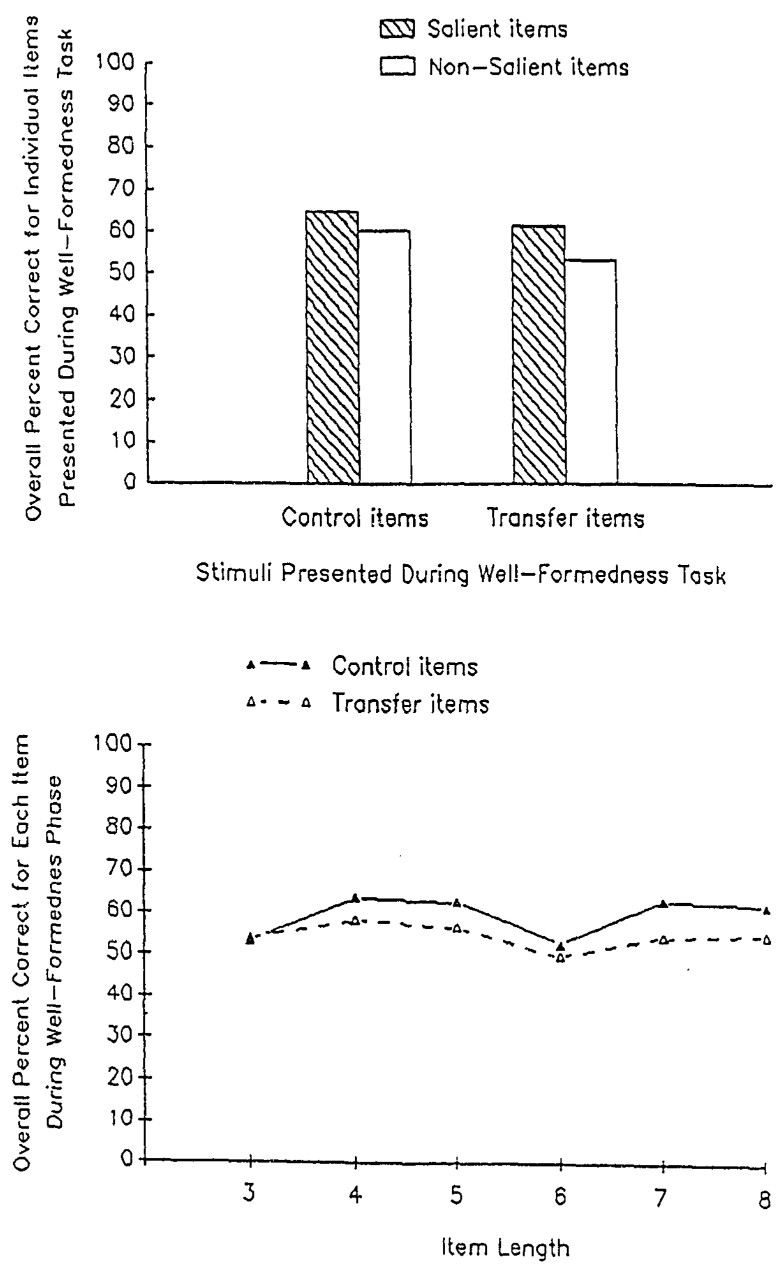


Figure 9. Overall percent correct for salient and non-salient items (top), & items ranging in length from 3-8 letters (bottom), from Experiment 2.

Experiment 3

If implicitly acquired information is represented in an exemplar-based, fragmentary, abstract, or combination of these forms, one area of investigation that could provide information in resolving this debate lies in determining how sensory information is processed in a nonconscious fashion. If subjects could be shown to transfer tacit knowledge across different sensory modalities, such a finding may support the abstractive position. However, if implicit processing were found to be limited to an initial input modality, then such knowledge would most likely be represented in a modality-specific form, lending support to the distributive and/or fragmentary view(s). The first study to investigate modality effects and implicit learning was conducted by Howard & Ballas (1982). As mentioned earlier, subjects participating in this experiment initially studied sequences of 'environmental sound' stimuli generated by an artificial grammar. For example, instead of the typical letter sequences (Reber, 1989), such as MMBQ, for example, subjects were presented with auditory stimuli such as "squeak"-"squeak"-"hiss"-"clang"; presentation occurred by subjects hearing the actual sounds or seeing the words on a computer monitor. When the task was changed from studying sequences to classifying new sequences as well-formed or not, subjects were found to perform equally well when presentation was either visual or auditory, and were able to transfer their acquired knowledge from the visual modality to the auditory modality. Howard and Ballas never ran a transfer from auditory to visual stimuli, but their results suggest that implicit processing is modality-free, and hence, most likely abstract.

Such modality nonspecificity of implicit processing was inferred to be a hallmark of nonconscious mental activity (Reber, 1989), but recent research on implicit memory has suggested otherwise, postulating that implicit processing may be limited to the visual modality (Schacter & Graf, 1989), "or, at best, tightly linked to the physical form of the input stimulus" (Schacter, 1992). In the typical cross-modal priming task, word-pairs, either alone (e.g., DOG-BICYCLE) or in the context of sentences (e.g., The DOG ran after the BICYCLE.), are originally presented in one sensory modality and then tested, by using some type of word-fragment completion task (where subjects are given a source word and a fragmented target word; e.g., DOG-B_ _Y_ _E), in another sensory modality.

By using this type of task, researchers have typically found that performance is poorer on the word-fragment completion task when the modality differs from the study presentation modality, as compared to when the modality remains the same for both study and test items (Jacoby & Dallas, 1981; Schacter & Graf, 1989; Roediger & Blaxton, 1987), although a few studies have reported instances of positive modality transfer (Kirsner, Milech, & Standen, 1983; Graf, Shimamura, & Squire, 1985). These latter results have been overshadowed in the literature by findings which have suggested that performance on an implicit priming task is greater when both the initial and test presentations of stimuli are visual, leading to the recent suggestion (Schacter & Graf, 1989) that implicit processing may be restricted to the visual modality. This hypothesis is questionable, however, as these researchers failed to test their subjects auditorally--the only transfer tested was from auditory to visual stimuli. In addition, as priming has been found to occur in both visual and auditory channels (Graf, Shimamura, & Squire, 1985), it may well be the case that although visual and auditory channels can process stimuli implicitly, the visual process might be better suited for nonconscious processing than the auditory process. This is quite plausible in light of the fact that the fragment completion task involves letters, which only exist in a visual form, whereas auditory stimuli (the spoken word) contain phonemes, not letters.

However, if nonconscious transfer is to occur, and if the knowledge acquired from implicit learning procedures is at least partially abstract then the explicit sensory nature of a stimulus should not be crucial to the acquistion or representation of tacit knowledge. Rather, one's internal representation of the structure of a stimulus domain is what should facilitate the representational process (Reder, Anderson, & Bjork, 1974). Accordingly, if transfer fails to occur in such a situation, it would most likely be due to a lack of appropriate attention to, and the subsequent misrepresentations of, structural relationships between stimuli or a largely specific, stimulus-bound form of mental representation.

In concordance with the current line of reasoning suggesting that implicit processes are modality-free, yielding abstract representations, Morton's (1969) original logogen model suggested that the sensory modality of the stimulus should not effect processing. Although this model dealt primarily with explicit processes and has not withstood experimental validation (Clarke & Morton, 1983), there are several aspects of it that may explain how sensory information becomes represented within the mind, a process which would facilitate transfer. The logogen model proposed that individual words are represented by a detector that is focused on various features of that word. After entering this 'detector system', the components of a word interact with the main cognitive system, which transforms the stimuli into salient information. After this transformation, the word then goes through the detector system once again, where it is formed into a response. Words that enter the system more often become 'common'; as a result, the detector can activate these words more easily than uncommon words. Where this model ties into the present discussion is that it is not modality specific; that is, the detectors are not limited to auditory or visual input--all the information concerning a stimulus is combined into one modality-independent representation. How a modified version of this system may explain the representational problem of implicit processes, and in turn the transfer of nonconscious knowledge, is as follows.

In an artificial grammar learning task, each individual stimulus may be viewed as a word, with entire stimuli coming together as sentences. When each word, and in turn, sentence, is perceived, it enters a transformational system where any sensory or modality-specific information is discarded in favor of the abstract qualities of the underlying rule structure of the 'sentence'. This rule structure then becomes salient, in a nonconscious fashion, allowing an individual to make grammaticality judgments when the rule is encountered later on. This system would explain transfer as resulting from the successful mapping of the rule structure, devoid of any particular sensory or symbolic information, from a stored representation to a novel, albeit transformed, stimulus. However, this line of reasoning is open to debate, as experimental evidence of nonconscious modality processing and transfer is still sparse.

For the scope of the present investigation, the aforementioned results of Howard & Ballas (1982), along with the model which is currently being developed, suggest that implicit processes should not be significantly different when the presentation modality is either visual, auditory, held constant, or shifted between study and test, as long as the type of processing and the associated representational system remain coordinate. The reasoning here is that "true" implicit induction is a process that yields a deep and abstract representation (Reber, 1989). Therefore, a superficial feature such as the modality of presentation should not effect nonconscious processing. However, the asymmetry in results between implicit learning and implicit memory modality transfer studies indicates that there are still questions that need to be answered in terms of the modality effect, transfer, and implicit processing. There are no compelling reasons, including those previously mentioned (Roediger & Blaxton, 1987; Schacter & Graf, 1989), why both inter- and intra-modality transfer should not take place implicitly, and that implicit learning should not be facilitated in either the visual or auditory modality. The reported findings of positive transfer of tacit knowledge (Reber, 1969; Howard & Ballas, 1982; Mathews, et al., 1989; Manza, 1992) suggest that implicit processes may be abstract and resistant to superficial surface manipulations. To add to the data base generated from the first two experiments, two additional studies were conducted to investigate the sensory nature of implicit processes and its relationship to implicit knowledge representation. One experiment addresses the structure of implicit representations by looking at transfer effects within and between the visual and auditory modalities, while the second study further delineates the sensory aspects involved in the acquisition, utilization, and representation of complex information. Building upon the incomplete results of Howard & Ballas (1982), and adding to the knowledge obtained from the first two experiments in the present investigation, it is expected that implicit learning will be equally robust in either the visual or auditory modality, and that shifting the modality between learning and testing will not effect performance adversely.

Method

Subjects, Design, and Materials

Sixty Brooklyn College students served as subjects, receiving either course credit or $7.50 for their participation. A 2 (visual or auditory learning) X 2 (visual or auditory testing) between subjects design yielded four fifteen-subject groups.

Stimuli presented during the course of the study, shown in Appendix D, were letter sequences generated from the finite-state artificial grammar shown in Figure 10. Although identical in structure

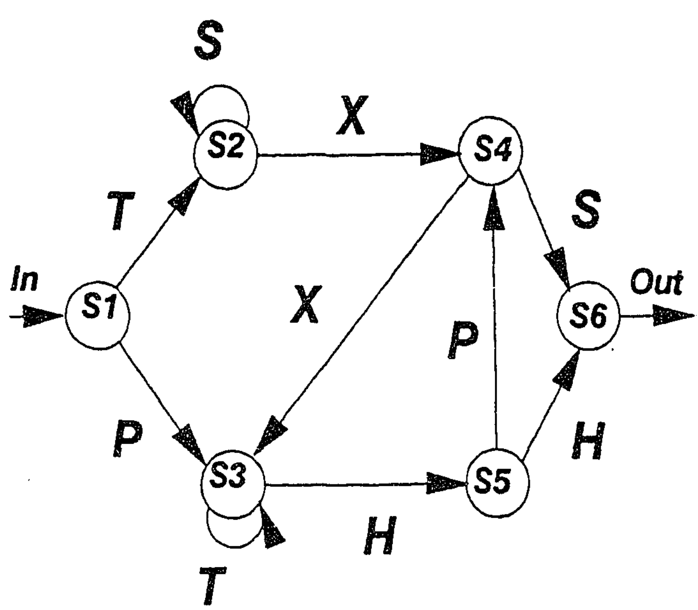


Figure 10. The artificial grammar used to generate stimuli for Experiment 3.

to the grammar utilized in Experiment 1, different letters were used to instantiate the stimuli. Twenty Grammatical strings were selected for presentation during the learning phase, with stimuli presented in the same order to all subjects. For the testing phase, fifty items, half of which were Grammatical and half of which were Nongrammatical, were selected for presentation. Each of the fifty strings was randomly presented twice, for a total of one hundred testing trials, with the items being presented to all subjects in the same order.

Procedure

Learning. Each of the 20 stimuli were individually presented, although the actual manner of presentation differed for visually and auditorally presented items. If presentation was visual, each item would appear at the center of a computer monitor for three seconds, after which the item would be removed and the subject would be prompted to reproduce the item. If presentation was auditory, stimuli were read aloud to each subject at the rate of two letters per second. After a three-second pause, the subject would repeat the string aloud and the experimenter would then record the response and verbally inform the subject whether their response was correct or incorrect. As in Experiment 1, the learning criterion was two consecutive correct reproductions.

Testing. The same instructions were administered as in the previous two experiments. The manner in which visual and auditory items were presented was identical to the learning phase, although subjects were allowed to make their decisions as fast as possible.

Subjects participating in auditory testing made their confidence ratings by looking at the scale on a sheet of 8.5" by 11" paper which remained in front of them during the entire phase. Following this phase, each subject was asked a set of questions, as in the previous two studies, designed to find out how much of the rule structure they were able to induce, including any specific rules they were using to make their well-formedness decisions.

Results and Discussion

Learning Phase

Trials to Criterion. An ANOVA found a significant main effect between the groups, $F(3,56)=6.0$, $MSe=1.4$, $p < .01$. A subsequent t-test showed that this difference resulted from the input modality having an effect on the acquisition process, as visual learning required significantly fewer trials than auditory learning to reach criterion (means of 3.1 vs. 4.4 trials to criterion, respectively), $t(58)=4.26$, $p < .0001$.

Testing Phase

Percent Correct. All four groups performed above chance in making grammaticality decisions, as evidenced by a significant Classical Normal Approximation to Binomial test on the total percent correct of 62.4, collapsed across the four groups, $z=18.9$, $p < .0001$. While an ANOVA found no significant overall group effect ($p=.19$), any interpretation of this finding would be premature, as the d' analysis, discussed later, revealed a contrary finding. As the d' measure is a more accurate relection of subjects' performance, any modality effects will be discussed in relation to the d' scores.

Confidence Ratings. An examination of the data on confidence ratings, presented in Table 9, yielded several important findings. Two repeated measures ANOVA's were initially conducted, the first looked at Group (1-4) X Accuracy (correct and incorrect decisions) and the second at Group X String Type (grammatical and nongrammatical). Overall, there was a significant difference between the groups, $F(3,56)=4.5$, $MSe=.5$, $p < .01$, plus the additional findings that subjects 1) were more confident when they made correct responses, although they were not explicitly informed as to the accuracy of their decisions, $F(1,56)=43.6$, $MSe=.04$, $p < .001$, and 2) were more confident with grammatical items over nongrammatical items, $F(1,56)=43.7$, $MSe=.06$, $p < .0001$.

To determine the cause of the significant Group effect and to see what effect the presentation modality or transfer manipulation had on subjects' confidence in their well-formedness decisions, 2 two-way ANOVA's comparing Manipulation (Transfer or Control) X Test Modality (Visual or Auditory) and Manipulation X Learning Modality (Visual or Auditory) revealed a main effect solely of Test Modality, $F(1,56)=12.55$, $MSe=6.52$, $p < .001$, with visual testing yielding higher confidence scores than auditory testing, an finding which accounts for the overall Group effect. Since the main effect of Manipulation failed to reach significance at the .05 level, the significant Test Modality effect is most likely due to the manner in which visual and auditory items were presented. When tested visually, each letter string

Table 9

Mean confidence ratings for each group during Experiment 3, across correct and incorrect judgments, & Grammatical and Nongrammatical items

	Accuracy		String Type		
Study-Test Modalities	Correct	Incorrect	G	NG	Means
V-V	5.40	5.10	5.36	5.14	5.25
A-A	4.52	4.31	4.56	4.27	4.42
V-A	4.76	4.44	4.82	4.38	4.60
A-V	5.15	5.03	5.21	4.97	5.09
Means	4.96	4.72	4.99	4.69	4.84

Note. V=visual presentation, A=auditory presentation.

remained on the screen until the subject made a decision, making each component of each visual string constantly available for sensory processing. However, when tested auditorally, once the string had been read, it was no longer available for immediate processing. This factor may have caused subjects to lose some of the structural qualities of each stimulus and therefore lead them to a less-confident rating.

Sensitivity Analysis. As in the previous two studies, a signal detection analysis revealed that, overall, responses were distributed normally. However, as can be seen in Table 10 and Figure 11, subjects who were presented with visual stimuli in both phases of the Experiment generated an overall higher d' score than the remaining groups, as there was a significant group effect ($F(3,51)=8.5$, $MSe=.02$, $p=.0001$ and Newman-Keuls test result, $p < .05$. Therefore, while visual processing may have a slight advantage in terms of implicit representations, nonconscious knowledge was obtained within the auditory modality as well. In addition, subjects were able to

Table 10

Mean d' and c scores for each group during the well-formedness phase of Experiment 3

Study-Test Modalities	Statistic	
	d'	c
V-V	.69	.08
A-A	.50	-.18
V-A	.48	.24
A-V	.48	.21

Note. A=auditory presentation, V=visual presentation

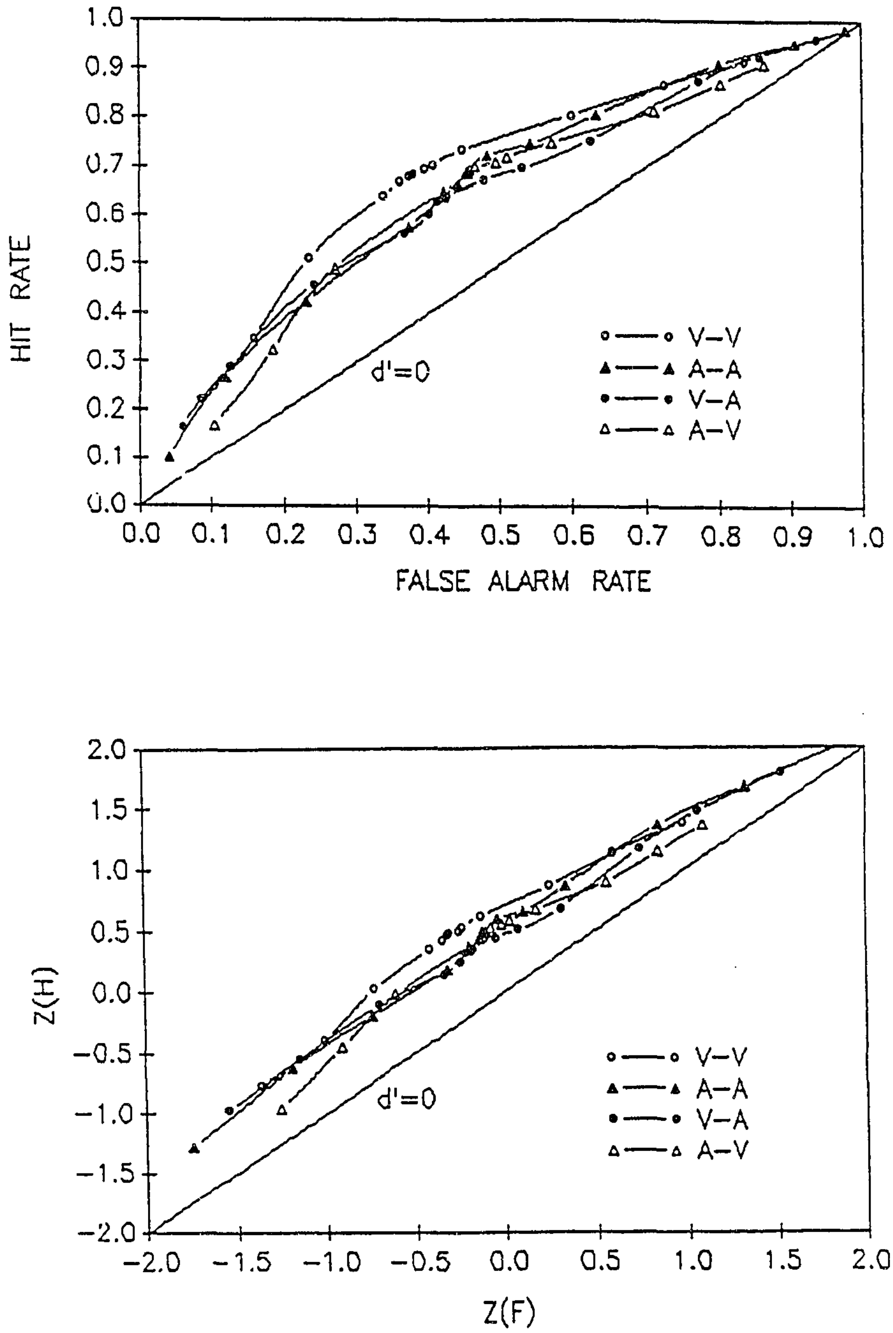


Figure 11. ROC and z curves for each group during Experiment 3.

transfer their tacit knowledge across both presentation modalities, suggesting that implicit systems are not entirely limited to the input form of stimuli. None of the obtained c scores revealed a significant response bias.

Consistency Scores. Subjects' consistency in responding to stimuli were again generated for each item by recording whether or not subjects responded correctly (C) or erroneously (E) to each item on each of its presentations. As in Table 11, while Transfer and Control subjects were found to have consistency scores which did not significantly differ from one another, $F < 1$, $p > .4$, there was evidence of the acquisition of non-representative knowledge in each group, as EE scores were significantly higher than the mean of the CE and EC scores in all groups except those receiving auditory stimuli during learning and testing ($z=1.61$, $p=.10$; all remaining z's > 2.48, p's $< .05$ (two-tailed).

Microanalysis of Responses. The first of the microanalysis data, displayed in Table 12, shows the now-familiar pattern of higher percentages correct for items with repeats in several locations per item. When these highly salient stimuli were compared to the remaining items, ANOVA's found no significant differences between the two item types in each condition, all F's < 1.2, lending additional support to the notion that tacit knowledge stores contain complex data are the results from the string length analysis. As displayed in Figure 12, subjects in all conditions responded with approximately equal accuracy to items of varying lengths, all F's < 1.

Table 11

Consistency scores for each group during Experiment 3

Study-Test Modalities	CC	CE	EC	EE	Total
	Consistency Scores				
V-V	51	18	11	20	71
A-A	45	16	18	21	66
V-A	44	16	17	23	67
A-V	47	16	14	23	70
Means	47	16	15	22	69

Note. A=Auditory presentation, V=visual presentation

Verbal reports. The answers that subjects provided to questions during the post-experiment interview revealed that the only rules which conformed to the structure of the grammar that subjects were able to verbalize were initial and terminal state rules. For example, most subjects reported that strings either "began with a P or a T" or "ended with an S". For the most part, this was the extent of correct verbalized rules. Few subjects were able to articulate the internal states of the grammar, and when they did, their 'rules' were

Table 12

Items with individual percentages correct in the top and bottom 10% of all test stimuli for Experiment 3

Study-Test Modalities	Proportion of Distribution Top 10% (Pct.)	 Bottom 10% (Pct.)
V-V	TXS (100)	PTTTTHPS (33)
	XXSHT* (93)	TXPH (30)
	PTHHH* (93)	TXHPS* (27)
	SHPXTHH* (90)	TSSSSXS (20)
	TXXTTHPS (90)	PTTTTTHH (7)
A-A	TXXTTHPS (93)	PHTHH* (37)
	XXSHT* (93)	PTTTHPHS* (37)
	TSXXTTHH (87)	PHTTTHH* (37)
	TSSXS (83)	TSXXPH* (33)
	TSSXXTHH (83)	PTTTTTHH (30)
V-A	TSXXTHPS (87)	PTTTHT* (30)
	SHPXTHH* (83)	PHPXTHPS (30)
	PHXPHXPX* (83)	TSSSSXS (27)
	HSTXHHS* (83)	PHTTTHH* (27)
	XXSHT* (83)	TXH (20)
A-V	XXSHT* (97)	TSSSSXS (30)
	PHPXHH (97)	TPTXS* (27)
	TSXXTHPS (90)	SXXHPS* (23)
	TSSXS (87)	PTTTTTHH (20)
	PTHHH* (83)	PHTHH* (13)

Note. *'s indicate a Nongrammatical item.

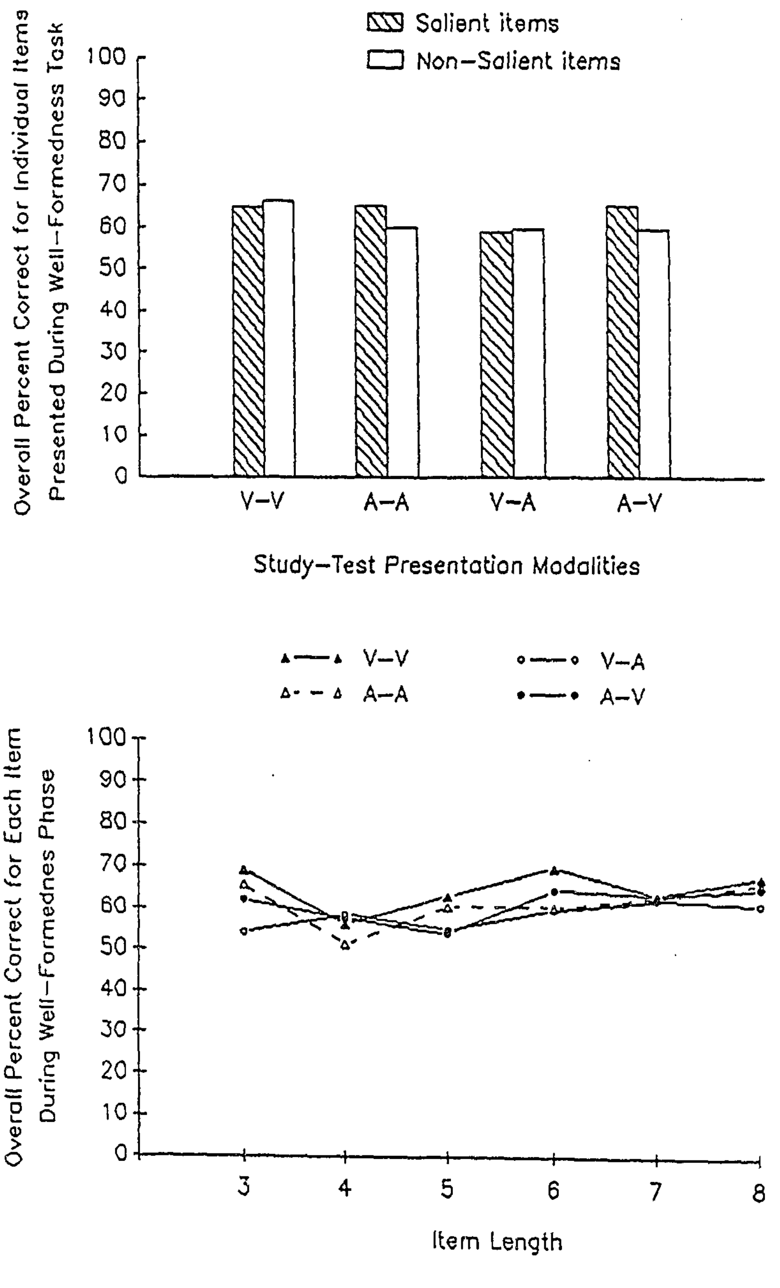


Figure 12. Mean percent correct for Salient vs. Nonsalient items (top), as well as items of length 3-8, for Experiment 3.

mostly inaccurate (e.g., "no letter could appear more than twice in a row"). Overriding any specific instantiations of the grammar, most subjects reported very vague ideas about rules they were using (e.g., "I looked for patterns"), and these statements lend additional support to the idea that individuals represent tacit knowledge in an abstract form.

Experiment 4

The results from Experiment 3, while showing that tacit knowledge can be acquired and manipulated equally well in either the visual or auditory presentation modalities, as well as lending support to the abstractive stance over the remaining two positions on implicit knowledge representation, did not clear up the representation issue altogether. The findings from this study were not terribly surprising. Subjects, when presented with a visual stimulus, could repeat it to themselves, creating an auditory version of the stimulus, or vice-versa for auditorally presented items. Such processing would logically facilitate cross-modal transfer. However, one way to eliminate this type of processing would be to remove the linguistic components of stimuli, making the above type of cross-modal processing more difficulty. In addition, one set of data that could clarify matters a bit more is reaction time (RT) to well-formedness judgments. Since the manner of presentation for the visual and auditory stimuli in Experiment 3 differed (i.e., for visual stimuli, the entire letter string remained on the screen until a well-formedness decision was made, while the components of the auditory stimuli were presented in a temporal fashion), RT data that would be comparable across conditions could not be collected. Therefore, to allow accurate measurement of decision latency, the temporal form in which visual and auditory stimuli were presented was made comparable for the final study, specifically, by instantiating the grammar as either physical locations or tones of varying pitches.

If nonconscious processing occurs abstractly, there should be no differences between Transfer and Control items, as the presentation modality would have no effect on processing. However, if implicitly acquired information is represented in a distributed or fragmentary form, Transfer RT's should be significantly slower than Control RT's. The logic here is that if tacit knowledge is represented in a modality-specific form, changing the modality of presentation would cause subjects to convert their representation from its original sensory form to a new sensory form, a process which would take longer than a process involved in an abstract system. Two additional reasons for altering the presentation form of stimuli were 1) to investigate the limits to which the surface form of stimuli can be altered, while still showing positive transfer, and 2) testing the hypothesized reason, discussed earlier, behind the higher confidence levels resulting from both visual learning and testing in Experiment 3.

Method

Subjects, Design, and Materials

Forty-eight Brooklyn College students enrolled in an introductory psychology course, and receiving course credit for their participation, were randomly assigned from a subject pool to partake in this study. The design was identical to that used in Experiment 3, with a 2 (visual or auditory learning) X 2 (visual or auditory testing) between subjects design yielding four 12-subject groups.

The grammar, in Figure 13, and items (see Appendix E) used in

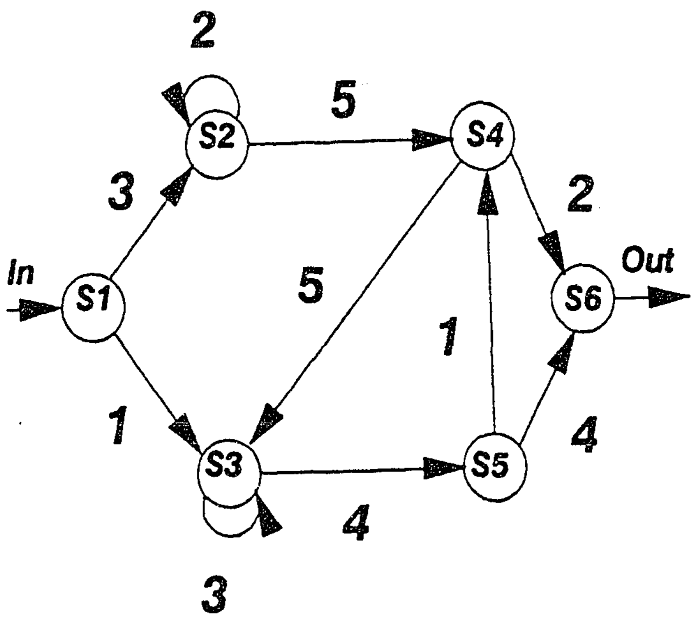


Figure 13. The artificial grammar used to generate stimuli for Experiment 4. For visual stimuli, the numbers 1-5 correspond to five boxes arranged from left (box 1) to right (box 5) on a computer monitor. For auditory stimuli, the numbers represent tones ranging from very low in pitch (tone 1) to very high in pitch (tone 5).

Experiments 1 and 3 was also used in this study, except that the letters which instantiated the grammar in those studies were replaced by stimuli that allowed for equal presentation times for visual and auditory items. For the current investigation, visual stimuli consisted of flashes of light that appeared for 250 msec each (followed by a 250 msec ISI) in one of five boxes arranged horizontally across the bottom 1/3 of a computer monitor. Auditory stimuli consisted of 5 tonal frequencies (190, 380, 750, 1500, and 3000 Hz) that were generated from the computer and sounded for 250 msec (with an ISI of 250 msec). In terms of how these stimuli correspond to the numbers in Figure 16, number 1 represents the leftmost box for visual stimuli and number 5 the rightmost box, while the lowest and highest pitched tones are represented by the numbers 1 and 5, respectively.

Procedure

Several changes in the procedure were necessary to accommodate the modified stimuli. First, during the learning phase, subjects could reproduce each sequence immediately after the final light or tone stimulus was presented by pressing the appropriate keys (# 1-5) on the keyboard. Also, when reproducing the flashes or tones, the appropriate stimuli, determined by the subject's response, would flash and/or sound on for 125 msec and subjects could not correct any errors. In addition, subjects with auditory learning went through a practice phase consisting of a random presentation of each of the five tones, individually, three times each, for a total of 15 practice trials. Each subject had to reproduce each individual tone correctly before hearing the next one. There was no such practice for visual learning and both visual and auditory testing. The well-formedness phase was also slightly altered. Subjects in Transfer conditions were first presented with a demonstration that allowed them to understand how the new stimuli (visual or auditory) corresponded to their learning stimuli. Briefly, subjects were shown that, for example, "box #1 corresponds to tone 1". At this point, box #1 would light up, followed by a presentation of tone 1. This demonstration continued for the remaining 4 stimuli, and subjects could repeat the entire demonstration again if they wished.

Results and Discussion

Learning Phase

Trials to Criterion. As in Experiment 3, the presentation modality had a significant effect on the number of trials to criterion. An ANOVA found a significant main effect of condition, $F(3,44)=16.9$, $MSe=16.1$, $p < .001$, and planned comparisons found this difference to be due to presentation modality; visual presentation (mean=3.2 trials) led to significantly fewer trials to criterion than auditory presentation (mean=11.4 trials), $F(1,44)=49.8$, $MSe=16.1$, $p < .001$.

Testing Phase

Percent Correct. The overall percent correct data was subjected to a Condition (1-4) X String Type (grammatical and nongrammatical) X Block (1 and 2) repeated measures ANOVA. Although there was a significant main effect of condition, $F(3,44)=4.6$, $MSe=106.7$, $p < .01$, as can be seen in Figure 14, all groups were found to be performing significantly above chance, $z=8.02$, $p < .001$. Interestingly, planned comparisons found that the condition effect was due to the fact that those subjects who were presented solely with visual stimuli during both acquisition and testing made their well-formedness judgments with better accuracy than the means of the

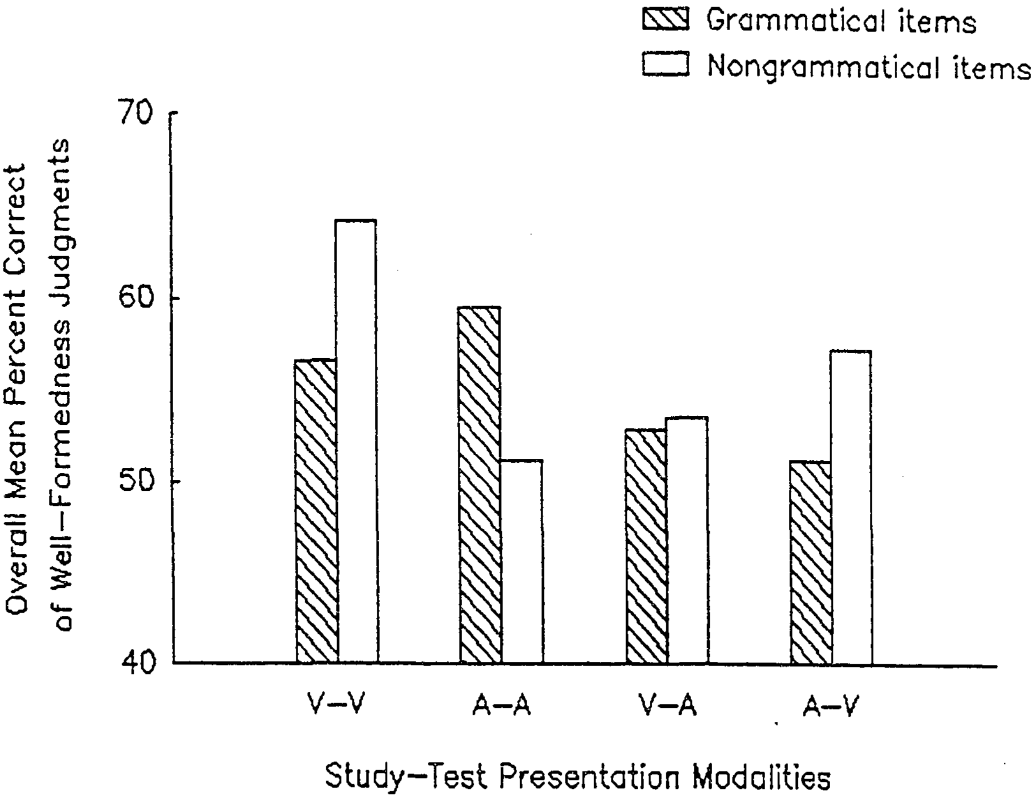


Figure 14. Mean percent correct of Grammatical and Nongrammatical items for each group, during well-formedness phase of Experiment 4.

remaining three groups ($F(1,44)=14.0$, $MSe=26.4$, $p < .001$). As with the sensitivity results from Experiment 3, it does seem that at least under these conditions that the visual modality is favored for implicit processing. The string type analysis revealed subjects performing equally well with grammatical and nongrammatical test items, ($F < 1$, $p > .6$), and all other main effects and interaction failed to reach significance at the .05 level.

Overall, therefore, in agreement with Howard & Ballas (1982) and contrary to the notion that implicit processing is modality specific (Schacter & Graf, 1989), tacit knowledge was acquired and utilized in both visual and auditory presentation modalities, lending some support to the hypothesis that tacit knowledge is at least partially represented in an abstract form. It may well be the case that subjects process the structure of items independent of the physical instantiations of stimuli. Such holistic representations (Wertheimer, 1945) would seemingly facilitate modality independent processing. Further support to this hypothesis was obtained by the analysis of reaction times to well-formedness decisions, discussed below.

Reaction Times. Investigating the latencies at which subjects make their well-formedness decisions can provide information in regards to how tacit knowledge is processed and stored. If complex information is represented in a distributive, modality-specific fashion, then there should be significant differences in the reaction times (RT's) between subjects who have their presentation modalities changed between learning and testing and those who don't. Specifically, a modality-specific representation would lead to longer response latencies when the presentation modality is changed between the 2 phases of the artificial grammar learning procedure, because one would have to convert the original sensory representation into one that corresponds with the new presentation modality. However, if one's representation is abstract and not limited to the input modality, there should be little or no effects of changing the presentation modality. To test these hypotheses, two repeated measure ANOVA's were carried out on the RT data; the first compared Condition X Block X String Type, while the second looked at Condition X Block X Judgment Accuracy (correct and incorrect).

There was no significant main effect of condition, $F=1.4$, $p > .3$, as subjects were able to make their well-formedness decisions equally fast when the presentation modality was held constant or shifted between study and testing (see Figure 15). In addition, even though subjects were never told about the accuracy of their decisions, they responded significantly faster when they judged items correctly (Mean=2.793 sec.), as compared to when they made incorrect decisions (Mean=2.962 sec.), $F(1,44)=5.0$, $MSe=.2$, $p < .05$. Latencies became faster as testing progressed, $F(1,44)=8.4$, $MSe=2.7$, $p < .01$, with mean response latencies for Blocks 1 and 2 being 3.207 and 2.550 seconds, respectively. This finding suggests that the more one's judgment is accessed, the more efficient processing becomes. Finally, RT's for G and NG items showed no significant differences between them, as the String Type main effect, also shown in Figure 15, failed to reach significance ($F < 1$).

Confidence Ratings. As mentioned earlier, since subjects are never informed as to the accuracy of their well-formedness decisions, the fact that grammatical items lead to slightly higher confidence scores than nongrammatical items (4.74 vs 4.60, respectively), $F(1,44)=12.5$, $MSe=.1$, $p < .01$. While the lack of a significant main effect of condition may indicate that subjects are equally confident

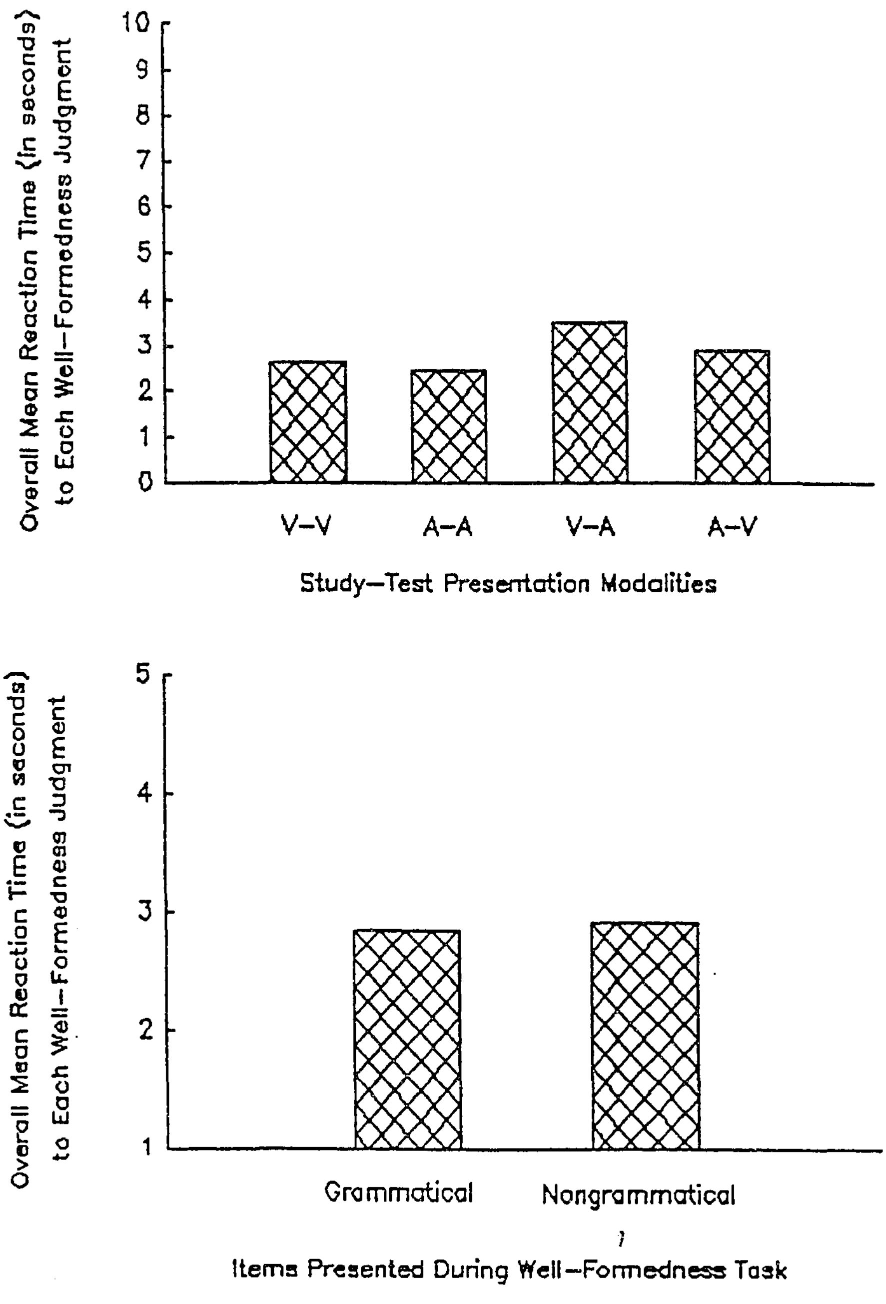


Figure 15. Mean reaction times (in seconds) to well-formedness judgments for each group (top), & Grammatical and Nongrammatical items (bottom), during Experiment 4.

confidence scores provide a good starting point for investigating nonconscious processes. Accordingly, subjects' confidence ratings were subjected to several analyses. Initially, an ANOVA on the factors of Condition X String Type X Block only found a significant effect of String Type, and as can be seen in Table 13, this was due to

Table 13

Mean confidence ratings for each group during Experiment 4, across Grammatical and Nongrammatical items & two test blocks

	String Type				
	Grammatical		Nongrammatical		
	Block				
Study-Test Modalities	1	2	1	2	Overall
V-V	5.24	5.19	5.15	5.03	5.15
A-A	4.73	4.54	4.46	4.44	4.54
V-A	4.76	4.46	4.44	4.50	4.54
A-V	4.55	4.49	4.52	4.25	4.45
Means	4.82	4.67	4.64	4.55	4.67

Note. A=auditory presentation, V=visual presentation when their presentation modality is held constant or shifted between learning and testing, there was a significant Condition X String Type X Block interaction ($F(3,44)=4.6$, $MSe=.04$, $p < .01$), which in light of the nonsignificant String Type X Block interaction ($F=1.1$, $p > .3$), appears to be due to the higher ratings given by group V-V subjects, as compared to the remaining three groups.

A secondary confidence analysis looked at the factors of Condition X Block X Accuracy, and yielded a significant main effect of Accuracy, with correct judgments (mean=4.74) leading to significantly higher confidence scores than incorrect judgments (mean=4.58), $F(1,44)=19.3$, $MSe=.1$, $p < .001$. This finding supports the identical conclusion from the first study, in that although subjects are never explicitly informed about the accuracy of their decisions, they are aware of something that is relevant to the task. What they don't know is the specific content of this implicit knowledge base. There was also a significant Accuracy X Block interaction ($F(1,44)=4.1$, $MSe=.1$, $p < .05$, which was the result of subjects reporting lower confidence ratings when correct as testing progressed (Block 1 mean=4.84, Block 2 mean=4.64) while their confidence with incorrect decisions remained fairly constant (Block 1 mean=4.60, Block 2 mean=4.57).

Sensitivity Analysis. From the ROC and z curves in Figure 16, it can be seen that, basically, the results from the previous studies were supported, as subjects responses were normally distributed. Also replicating the findings from Experiment 3, subjects in group V-V had higher d' scores (see Table 14) as compared to the remaining

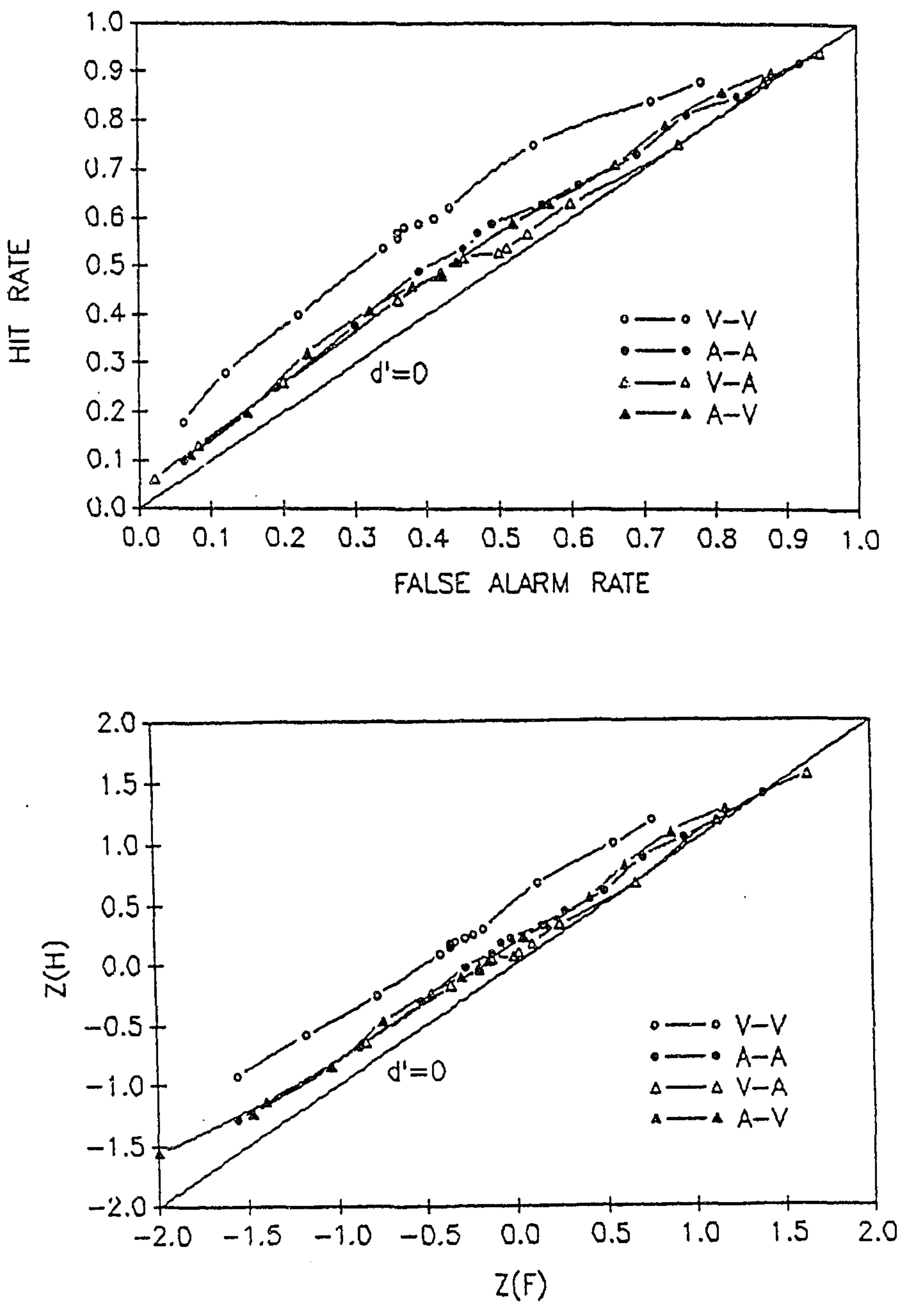


Figure 16. ROC and z curves from Experiment 4.

Table 14

Mean d' and c scores for each group during the well-formedness phase of Experiment 4

Study-Test Modalities	Statistic d'	c
V-V	.52	.05
A-A	.19	.11
V-A	.14	0
A-A	.19	-.05

Note. A=auditory presentation, V=visual presentation.

three groups, as evidenced by a significant group effect, $F(3,51)=48.9$, $MSe=.03$, $p < .001$, and Newman-Keuls test ($p < .05$). Subjects in the pure auditory condition (A-A) showed a slight bias towards responding 'No'.

Consistency Scores. To analyze the degree to which subjects' responses are consistent, and to determine if any non-representative knowledge was obtained, consistency scores were once again subjected to two separate analyses. Initially, an ANOVA was carried out on subjects' overall consistency in responding, derived by summing each subject's CC and EE score. As can be seen in Table 15, and replicating the results from Experiment 3, no significant difference

Table 15

Consistency scores for each group during Experiment 4

Study-Test Modalities	Consistency Scores CC	CE	EC	EE	Total
V-V	42	20	17	21	63
A-A	38	18	17	27	65
V-A	35	19	17	29	64
A-V	38	16	17	29	67
Means	38	18	17	27	65

Note. A=auditory presentation, V=visual presentation

was found to exist between the groups, $F < 1$. However, as in the previous three studies, there was evidence of the formation of nonrepresentative rules, as the EE rate was significantly higher than the mean of the CE and EC score in all groups except those subjects receiving visual testing for both phases of the study (all z's > 3.5, all p's < .01). Finally, the effects of Learning Modality, Testing Modality, and Modality Transfer were examined by conducting several

2 way ANOVA's, with the results uncovering no significant main effects or interactions existing around these variables.

Microanalysis of Responses. The overall percent correct scores from each individual item was subjected to the same analyses undertaken in Experiments 1-3, with the results mimicking those findings. As seen in Table 16, subjects once again were more accurate with stimuli that had repetitions in different places. However, these salient items were not the entire basis of subjects' knowledge systems, as the ANOVA comparing salient and non-salient items (the data from which are in Figure 17) in each group failed to find any significant differences (all F's$(1,49) < 3.6$). Finally, the string length analysis confirmed previous null findings, as no differences were found to exist between items of various lengths (see Figure 17; all F's < 1).

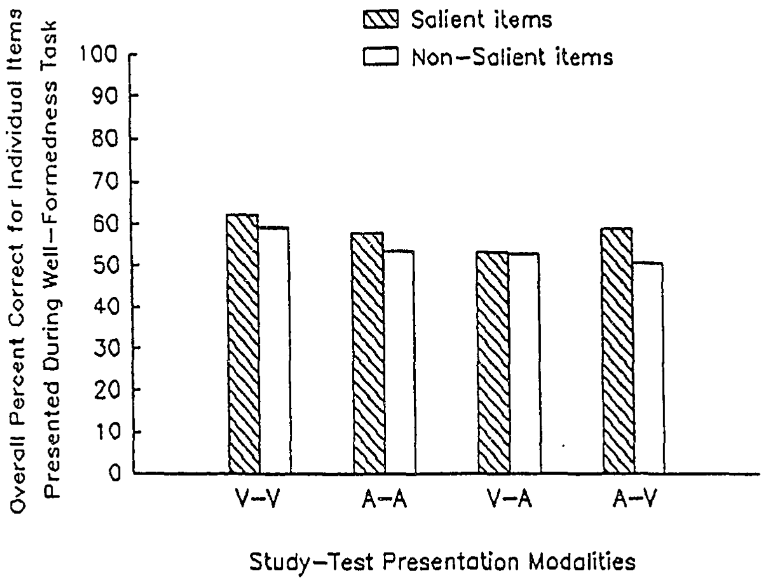
Verbal Reports. In terms of the inductive processes involved in acquiring the rules of the AG, subjects were once again unable to verbalize any substantiative rules of the grammar other than the starting and finishing states and general pattern rules (e.g., "There could be two identical tones/flashes in a row, but not three."). In addition, when subjects did report specific letter/tone/flash patterns, they were predominantly rules that were non-representative of the structure of the grammar, a finding which once again begs to the issue that subjects develop personal grammars, parts of which may be accurate and parts of which may not.

Table 16

Items with individual percentages correct in the top and bottom 10% of all test stimuli for Experiment 4

Study-Test Modalities	Proportion of Distribution: Top 10% (Pct.)	Bottom 10% (Pct.)
V-V	35533412 (91)	35545* (36)
	134444* (86)	3553413* (32)
	31352* (82)	13333412 (32)
	13441112* (82)	144 (14)
	14344* (77)	13333344 (9)
A-A	32553412 (91)	13334142* (33)
	13333412 (86)	352 (33)
	3514* (83)	3553413* (33)
	355344 (83)	32255422* (25)
	31352* (83)	13333344 (17)
V-A	3344* (75)	13333344 (33)
	133344 (71)	13441112* (29)
	32252 (71)	133412 (25)
	35545* (71)	32255422* (21)
	31352 (71)	13334142* (21)
A-V	55243* (79)	3553413* (33)
	14153412 (79)	3222252 (29)
	14153344 (79)	13333344 (25)
	14333* (75)	15145433* (25)
	354* (75)	352 (17)

Note. *'s indicate a Nongrammatical item.



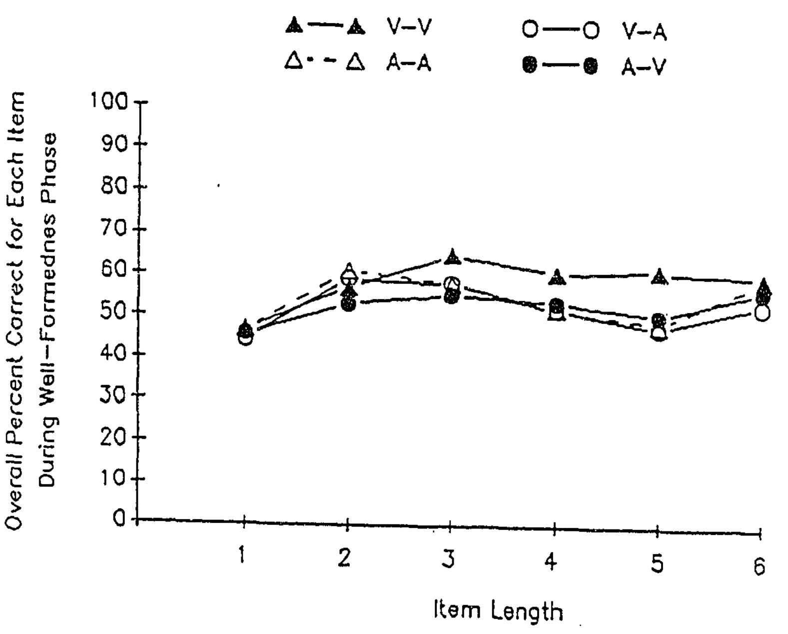


Figure 17. Mean percent correct for Salient vs. Non-salient items & items of length 3-8 for Experiment 4.

General Discussion

The original aim for conducting the preceding studies was to provide insights into the form of the mental representation of tacit knowledge. The primary schools of thought on this matter are that nonconscious information is represented in a specific, exemplar based fashion (Brooks, 1978), a fragmentary form (Perruchet & Pacteau, 1990), or in an abstract, prototypical form (Reber, 1989; Mathews, 1991). In order to determine which of these positions captures the data most accurately, four transfer studies were conducted. The logic for implementing a transfer procedure to study knowledge representation was that since the learning and testing stimuli, although physically different, shared structural commonalities, the occurrence of positive transfer would provide insight into how knowledge acquired primarily outside of conscious awareness was stored. The results obtained from each of the four current studies revealed evidence of positive transfer across orthographic forms or sensory modalities, findings which can be applied to a theoretical framework of the implicit learning process.

Several such frameworks currently exist in the literature, but as cited earlier, while each model contains mechanisms that can seemingly model some aspect of implicit learning, they have all failed with respects to one crucial point of interest: transfer. To date, not a single IL system has been able to adequately capture the transfer of nonconscious knowledge, and this may be due to the limitations of the individual systems. One of the main points of debate outlined earlier is that each of these models is limited to the specific form in which data are input into the system; that is, none of the existing models has been shown to be capable of abstraction. From the present results, as well as from recently published work (Vokey & Brooks, 1992), the form in which tacit knowledge assumes in the mind seems to be one of a synergistic relationship between both specific and abstract components. An accurate model of IL should therefore be capable of exploiting this relationship. It would perhaps be more beneficial to attempt to combine several aspects of existing models into a new theoretical framework that can accurately describe the implicit learning process. Before this is undertaken, however, a brief review of the current state of implicit learning is in order.

Servan-Schreiber & Anderson's (1990) competitive chunking model hypothesizes that a chunking mechanism, available in long-term memory, organizes chunks of artificial grammar-generated stimuli in a hierarchical fashion, with simple chunks at the bottom and complex chunks at the top of the hierarchy. When one encounters a stimulus, this mechanism begins processing the information from the simplest level available, working up the hierarchy. As chunks are accumulated in memory, subsequent stimuli are compared to stored exemplars to determine if a stimulus is well-formed according to the rules of the grammar. The drawback to this system is that it is based on concrete exemplars. This model fails to explain how one would transfer the information about the chunks to a new physical instantiation, which is what happens when transfer occurs. Another model which falls into the trap of being unable to handle transfer is Cleeremans & McClelland's (1991) simple recurrent network. However, this processor is not meant to model the type of task employed in the present experiments; the heart of this model is one that is capable of predicting and generating covariational patterns, so its relevance to the present discussion is unfounded.

However, another model, one which tries to handle the abstraction process, is Roussel, et al.'s (1990) THIYOS model. This system postulates that rules of grammaticality are stored in a production system, with statements of the sort "IF [a certain condition is met], THEN [perform an action]". The device that Mathews (1991) has argued for being capable of handling the abstraction process is what is known as the forgetting algorithm. This mechanism retains some information about a stimulus, and 'forgets' other information. Presumably, the important aspects of a stimulus are retained, while less vital data are forgotten. As an example, if the string VVTRDV were to be encountered, the forgetting algorithm might remove the VVT chunk to create the rule "IF a sequence ends in RDV, THEN it is a valid rule". What makes this rule abstract, according to Mathews (1991), is that since the specific beginning of the string is eliminated, the new rule is more compact and can account for more instances of the grammar, one only needs general knowledge of the existence of the forgotten component in order to make grammaticality judgments. Although this is a good attempt at explaining the abstraction process, it too cannot handle the transfer process.

A type of model that might be able to accurately model the implicit transfer process is one similar to a model suggested recently by Nosofsky, Kruschke, & McKinley (1992). This system is a connectionist network that uses exemplar-based data as well as adaptive learning rules to represent knowledge. It first inputs data through several input nodes, after which information is sent to several 'hidden nodes', which represent knowledge in an exemplar-based form. From there, items are given a learned association weight and then output from the system. Although not a model of implicit learning, the theoretical assumptions this approach makes can be considered in developing a theoretical, and eventually practical, model of the implicit learning process. However, this model, while containing learning mechanisms, is still bound to represent knowledge according to its input form. This would present a problem for transfer, in that some type of abstraction device seems necessary in order for transfer to occur.

Toward a model of implicit learning

Although each of the preceding approaches are somewhat viable in explaining how nonconscious knowledge is processed, no current model can adequately handle true transfer. In addition, these models handle general and specific knowledge in different manners. The results from the current experiments suggest that implicit knowledge is represented by an eclectic mix of fragmentary, exemplar-based, and abstract components, a hypothesis arrived at by looking at the implicit learning process itself, as well as the models proposed in the psychological literature.

When subjects initially learn the structure of the grammar, they do so by memorizing or otherwise coding the explicit sequence of symbols of the grammar. Therefore, tacit knowledge is most likely initially represented in a chunked form of fragmentary and distributive information, much like the competitive chunking model of Servan-Schreiber and Anderson (1990). However, as the load on the memorial system that stores these chunks of knowledge increases, some of these instance-based elements may come to be replaced by more efficient abstract components. This proposal falls in line with Mathews' (1991) 'forgetting algorithm', a procedure which retains some tacit knowledge yet discards other pieces of nonconscious data.

For example, assume for a moment that a subject is trying to keep the following Grammatical items stored in memory during the acquisition portion of an IL procedure: HXHTRRX, HXTRR, HXHRHXHT, HXHRHHT, and HXHRM. Although these items are complex, the one underlying commonality they share is that they all begin with the bigram HX. The forgetting algorithm, in an attempt to maintain a streamlined memorial system, may choose to store only this bigram in its raw form, while the remaining portions of these items are 'forgotten' in the sense that the explicit symbols are replaced by some abstract rule such as "middle portions of strings may contain single occurrences of symbol types 1, 2, and 3". While this eclectic 'rule' may not capture every stimulus within this domain, it allows processing to become more efficient, and in turn (hopefully), more accurate. In a broader sense, the elements that remain instantiated are most likely the highly salient initial and terminal states of the grammar, while the inner, more complex states become represented abstractly. Verbal reports gathered from subjects as they were learning grammar generated items (see Mathews, et al., 1989), validate this hypothesis.

With the form in which initial learning may occur in place, the discussion can now turn to how the transfer process itself might occur, something which, mentioned above Competitive Chunking and THIYOS have been incapable of capturing. Any theoretical model needs empirical evidence for support, and the data to possibly explain the implicit transfer process were gathered in the present set of experiments by measuring subjects' reaction times to their well-formedness decisions.

Reaction times were collected in 3 of the 4 studies conducted, with the results from two of those three experiments showing reaction times to Transfer items not differing significantly from reaction times to Control stimuli. In addition, the data from the remaining study (Experiment 1) were in this direction, as a Block effect showed Transfer and Control RT's eventually becoming closer to one another. These findings have bearing on the representation of implicit information due to the following line of thought. Some researchers hold that tacit knowledge is represented in its specific, raw form (Brooks & Vokey, 1991); transfer of such knowledge is held to occur by subjects drawing analogies from one surface domain to another. On the other hand, others believe that tacit knowledge is represented abstractly, allowing transfer to occur by simply mapping a prototype to any surface form that data might be presented in (Mathews, et al., 1989; Mathews, 1990; Reber, 1990). The former, analogical, transfer process would seemingly take longer than the latter process, since making analogies between different surface forms would apparently take longer than mapping a prototype. Therefore, with data from 2 out of three studies showing Transfer RT's about equal with Control RT's, the abstractive position seems to have more support. This statement, however, in no way suggests that abstract elements are the sole foundation of nonconscious representations. The point throughout these studies has been that tacit knowledge is represented by a mixture of distributive and abstract components. It may be the case that the exemplar-based elements play a more vital role at a lower, initial level of processing, while the more general, abstract elements moderate the higher level processes such as mapping stimuli across different sensory and orthographic domains. Although Brooks (Brooks & Vokey, 1991; Vokey & Brooks, 1992) has suggested that exemplar-based factors play a role in the higher-level mapping of one stimulus form onto another by way of 'abstract analogies', he has not offered empirical support for this claim. This fact, combined with the finding from Experiment 2 that subjects could not match Transfer symbols to their appropriate Control symbols above chance performance, seems to suggest that abstract elements direct the transfer process, while stimulus-specific components are more responsible for moderating the initial acquisition of complex information.

An additional set of data that seemingly supports the notion that tacit knowledge is primarily represented by abstract components centers on the modality-nonspecificity findings of Experiments 3 and 4. That tacit knowledge can be acquired equally well in either the visual or auditory modalities confirms and extends the results of Howard & Ballas (1982), in that nonconscious learning processes are not limited to any sensory modality. These studies suggest, contrary to other opinions (Roediger & Blaxton, 1987; Schacter & Graf, 1989), that implicit processing may be modality independent and, therefore, mostly abstract. On a broader level, the current results seem to indicate that implicit processing is not dependent on the explicit nature of the stimuli; it is the underlying structure that is crucial to processing. Just how this structure might be processed in order for a eclectic representation (consisting of fragmentary, distributive, and abstract components) to be formed is best illustrated by an example.

Assume that a subject is initially presented with the stimulus BMMMQBH. At an initial level of processing, each individual letter would have to have its sensory components understood. Once this was accomplished, any irrelevant sensory data could be stripped away, leaving the letters in their raw form. At this point, some type of chunking mechanism, similar to the one proposed by Servan-Schreiber and Anderson (1990), would take over, breaking the stimulus into several chunks of knowledge (e.g., B-MMM-QBH). As this item, and items similar to it, come to be encountered on several occasions, the chunking becomes more efficient, representing the item in as few chunks as possible, with the final whole item being stored as 1 chunk (e.g., BMMM-QBH ---> BMMMQBH). At this point, the stimulus has gone from an initial fragmentary representation to a more exemplar-based representation. Once this has been accomplished, an 'abstractor' would come into play, removing irrelevant surface data, leaving only structural elements in the representation (e.g., BMMMQBH ---> B-(S2)(S2)(S2)-QBH ---> B (S2)(S2)(S2)(S3)(S4)(S5) ---> (S1)(S2)(S2)(S2)(S3)(S4)(S5)).

Such a system would be highly efficient, and would result in Transfer judgment accuracy being equal to Control judgment accuracy. However, this was not the case. In all but Experiment 3, Transfer accuracy lagged significantly behind Control accuracy. Therefore, the above proposed representation is not an accurate reflection of the stimulus environment. A more accurate picture would be one in which each step in the above process occupies a certain level in a representational system, in the following fashion:

(S4)(S2)(S2)(S2)(S3)(S4)(S5)

B--(S2)(S2)(S2)(S3)(S4)(S5)

B--(S2)(S2)(S2)--QBH

BMMMQBH

BMMM--QBH

B--MMM--QBH

B--M--M--M--Q--B--H

Processing in this system would begin at the bottom level and attempt to proceed to the highest level. The reason the term "attempt" is used is that some knowledge would likely not reach the top level of the system. As a result, that knowledge may remain represented at its last level of processing, an occurrence accounting for the poorer Transfer accuracy as compared to Control accuracy. Such a system, therefore, would obviously contain a mixture of fragmentary, exemplar-based, and abstract components. This assumption is somewhat supported by the finding from each of the preceding four studies that highly salient items were responded to with approximately equal levels of accuracy when compared to non-salient items. Subjects are apparently able to store both specific instances as well as complex, abstract information within the same tacit system. The only question that remains, therefore, is which form of knowledge, if any, has the greatest influence on this system. The answer to this question lies in the reaction time data, as outlined above. If the fragmentary and/or exemplar based components had a greater influence than the abstract components, then RT's for Transfer items should lag behind those for Control items, since Transfer stimuli would have to be converted into a exemplar-based form which coincided with the stored representation. On the other hand, if abstract elements assumed a greater role in the tacit knowledge representational system, this RT difference would not exist, due to the fact that no conversion would be necessary. To recapitulate, 2 of the 3 studies in which RT's were collected found no significant differences between Transfer and Control RT's, a finding which supports the idea that although implicitly acquired information may be represented in a variety of formats, the primary component of a tacit knowledge representation appears to be abstract, prototypical knowledge about the stimulus environment.

Conclusions

The issue of knowledge representation, independent of conscious or nonconscious elements, is an exceedingly complex matter. Transfer procedures have come to be a fairly powerful measure of any theory of knowledge representation, this being the case especially in the field of implicit learning. How implicitly acquired knowledge is represented has come to be investigated only very recently. The results obtained from the present studies are in line with current proposals (Vokey & Brooks, 1992) suggesting that tacit knowledge is stored eclectically via both distributive and abstract elements. However, there is still a need for a model which can accurately reflect and predict nonconscious functioning; although several exist, they are, as mentioned, inadequate. In addition, the results of Manza (1992) suggest that the development of implicit representations can be influenced by the type of processing that occurs during the initial acquisition of the rule structure of an artificial grammar; this issue, too, needs to be explored before any firm and definitive conclusions about the representation of nonconscious knowledge can be made. The theoretical model outlined above is a possible starting point for clarifying these matters, although it is not clear if such a model: (a) could be instantiated, and (b) could capture the transfer of nonconscious knowledge.

To conclude, one of the most important aspects of nonconscious processing of complex information is the fact that it seems to hinge upon one's ability to extract the deep structure of the stimulus environment (Reber, 1989) rather than the surface elements in order to garner an adequate understanding of the stimulus domain. In order to accomplish this, processing must be centered on relevant data from the stimulus display. In the transfer setting, this means concentrating on the structural components of stimuli and not their surface features. This can be done by utilizing both exemplar-based and abstract processing mechanisms, although the roles of these two processes in the representation of tacit knowledge have yet to be made totally clear.

The present results offer encouragement for delineating the implicit representation process, however. As mentioned above, the reaction time data collected in three of the four investigations suggest that stimulus-specific data control the initial lower levels of nonconscious processing, while abstract elements are formed by stripping down those distributive data at a higher level of functioning. Such a synergistic relationship between memorial items is seemingly necessary to fully capture the implicit learning process, for previous work, as well as the current studies, has indicated that studying nonconscious processes is an extremely complex task; the best way of handling these matters is to utilize the best aspects of the varying theories currently in existence in relation to the acquisition and representation of complex information.

Appendix A

Questions asked to subjects at the conclusion of each session (with sample answers).

1. During the learning phase, how were you memorizing the items presented to you?

   "Repeated the letters to myself"
   "Broke them down into groups of 2 or 3"
   "Tried to make words out of them"

2. In general, what were you basing your Yes/No judgments on for the testing phase?

   "I don't know"
   "I looked for patterns"
   "Familiar letters"

3. Were there any specific rules you remember looking for and/or using?

   "Nothing specific--just a vague sense"
   "Letters (flashes) (tones) couldn't repeat more than twice in a row"
   "[Specific letters, flashes, tones] at the beginning and others at the end"

4. If an item didn't seem to follow any rule you were using, what made you decide Yes or No for that item?

   "An intuition for that item"
   "I guessed"

Appendix B

Stimuli presented to subjects for Experiment 1

Learning Phase Stimuli

Symbol Set A	Symbol Set B
MHXH	LJZJ
MXXMQBH	LZZLRWJ
MHXXMQBH	LJZZLRWJ
MXH	LZJ
BQBXMQBH	WRWZLRWJ
BMQQ	WLRR
MXXMMQBH	LZZLLRWJ
BQQ	WRR
MHHHXXQQ	LJJJZZRR
MHHXXQBH	LJJZZRWJ
MHXXMQQ	LJZZLRR
BMQBXQQ	WLRWZRR
BQBXQQ	WRWZRR
MXXQBXQQ	LZZRWZRR
MXXMMMQQ	LZZLLLRR
MHHXXQQ	LJJZZRR
MHHXH	LJJZJ
BMMMQBH	WLLLRWJ
BMQBH	WLRWJ
BQBXQBH	WRWZRWJ

(Appendix B continues)

Appendix B (continued)

Stimuli presented to subjects for Experiment 1

Testing Phase Stimuli (* indicates a Nongrammatical item)

Set A	Set A (mix)	Set B	Set B (mix)
MXXMMQBH	QHHQQXMB	LZZLLRWJ	RJJRRZLW
*MHXXBQ	*QBHHMX	*LJZZWR	*RWJJLZ
*BQMQQ	*MXQXX	*WRLRR	*LZRZZ
BQBH	MXMB	WRWJ	LZLW
*MXBQ	*QHMX	*LZWR	*RJLZ
*BHXH	*MBHB	*WJZJ	*LWJW
*BXBQXQMM	*MHMXHXQQ	*WZWRZRLL	*LJLZJZRR
MHXXMQBH	QBHHQXMB	LJZZLRWJ	RWJJRZLW
*XXHQM	*HHBXQ	*ZZJRL	*JJWZR
*BQXBQXBX	*MXHMXHMH	*WRZWRZWZ	*LZJLZJLJ
MHXXMMQQ	QBHHQQXX	LJZZLLRR	RWJJRRZZ
BMMMQQ	MQQQXX	WLLLRR	LRRRZZ
*HQBXMQQ	*BXMHQXX	*JRWZLRR	*WZLJRZZ
MHXXQBH	QBHHXMB	LJZZRWJ	RWJJZLW
BQBXMMQQ	MXMHQQXX	WRWZLLRR	LZLJRRZZ
BMMMMQBH	MQQQQXMB	WLLLLRWJ	LRRRRZLW
*BMQBXQHB	*MQXMHXBM	*WLRWZRJW	*LRZLJZWL
*QHMXQQH	*XBQHXXB	*RJLZRRJ	*ZWRJZZW
*BMQQQQ	*MQXXXX	*WLRRRR	*LRZZZZ
*BMMQBQH	*MQQXMXB	*WLLRWRJ	*LRRZLZW
*MXQ	*QHX	*LZR	*RJZ
*MBMXH	*QMQHB	*LWLZJ	*RLRJW
BQBXQQ	MXMHXX	WRWZRR	LZLJZZ
*BMQBBBH	*MQXMMMB	*WLRWWWJ	*LRZLLLW
BMMQBXQQ	MQQXMHXX	WLLRWZRR	LRRZLJZZ
*MXXQX	*QBBHB	*LZZRZ	*RJJZJ
BMMMMMQQ	MQQQQQXX	WLLLLLRR	LRRRRRZZ
*HXXQBH	*BHHXMB	*JZZRWJ	*WJJZLW
MBQQ	QMXX	LWRR	RLZZ
*QBXMQQ	*XMHQXX	*RWZLRR	*ZLJRZZ
MXH	QHB	LZJ	RJW
MXXMMQQ	QHHQQXX	LZZLLRR	RJJRRZZ
MHHHHXH	QBBBBHB	LJJJJZJ	RWWWWJW
*BQMMM	*MXQQQ	*WRLLL	*LZRRR
*MXQBH	*QHXMB	*LZRWJ	*RJZLW
MXXMQQ	QHHQXX	LZZLRR	RJJRZZ
BQBXMQBH	MXMHQXMB	WRWZLRWJ	LZLJRZLW

(Appendix B continues)

Appendix B (continued)

Stimuli presented to subjects for Experiment 1

Testing Phase Stimuli (* indicates a Nongrammatical item)

Set A	Set A (mix)	Set B	Set B (mix)
*MMQQ	*QQXX	*LLRR	*RRZZ
MHHXXQBH	QBBHHXMB	LJJZZRWJ	RWWJJZLW
*BMMMQM	*MQQQXQ	*WLLLRL	*LRRRZR
BMQBXQBH	MQXMHXMB	WLRWZRWJ	LRZLJZLW
*MHHXXQHH	*QBBHHXBB	*LJJZZRJJ	*RWWJJZWW
BQQ	MXX	WRR	LZZ
MHHXXMQQ	QBBHHQXX	LJJZZLRR	RWWJJRZZ
*MXXMQBM	*QHHQXMQ	*LZZLRWL	*RJJRZLR
*BMMBH	*MQQMB	*WLLWJ	*LRRLW
BMMQBH	MQQXMB	WLLRWJ	LRRZLW
MXXQQ	QHHXX	LZZRR	RJJZZ
MHXXQQ	QBHHXX	LJZZRR	RWJJZZ
MHHXH	QBBHB	LJJZJ	RWWJW

Appendix C

Stimuli presented to subjects for Experiment 2

Learning Items

MHXRHMRH	MXRHHHM
HXHT	HMRTTMXR
MHXRMHR	MXRTTMXR
MXRTMXR	HMTRRRRX
MHXRHH	HXHTRX
MHXRHXHT	MXTRX
HMTRRRR	HMRTHMT
HXHRM	MHRHHHH
MHXRTHMT	HXHRHXTR
HXHRHMRH	MXRTTHMT
HXHRHXHR	HMRHH
MHXRMHXT	MXRHMTRR
HXHRTHMT	MXTRR
HXTR	HMRHXTRR
MXR	HMRHMTR
HMRMXTX	HXHRMHXT
HXHRTMXR	MHXRMXRH
MXRMXT	MXRMXRHM
HMRTHMRM	MHXRTMHR
MHXTRR	HMRHHHHM
MXTRRRX	MHRM
HMRHMT	HXHRMXT
MXRTHMTX	HMRMXR
HMRTMXR	MXRTHXT
HXHRHMR	MHXRTMXT

(Appendix C continues)

Appendix C (continued)

Stimuli presented to subjects for Experiment 2

Testing Phase Stimuli (* indicates Nongrammatical item)

Transfer Items		Control Items	
*ZJZBZJBZ	*BZJFBBBB	HXHRM	*HXHTMXR
*QZJQB	ZQB	*MTXTRR	MHXRMHRH
*QJBFFQJJ	ZJZBZJZF	*HXHMMHRH	MHRHHHHM
*FQB	QJBZ	HXHRMHXR	MXRHXTRR
ZJFBB	ZQBFFQJF	*MRXRMHRM	*HXMRHHM
ZJZBZQBZ	QJBFZJF	MXRMXTRR	MXRHMRHM
*ZJZQQZBZ	QZBZZZZQ	*MHRRMXR	HMTRRRRX
*QJBQFBZ	ZJZBFZJF	HXHTRRX	*MXRMTRH
*QJFF	*ZQBFBJF	HXHRHMRH	*HMRMRT
*ZJQBZZQ	QJBQJFB	HXTRR	*MXRXMRM
QZJBZJF	QZJBZQB	HMRHHHHH	MXR
ZQBZQ	*QJBFZZZQ	*HHXRTMXR	*RHXTRRRR
*ZZZBQJBZ	ZQFBBBBJ	MXRTMXTR	MXRTTMXR
*ZBZF	QZJBFZJF	HXHRHXHT	MXRTHMT
*QFBZZQ	QJBFZQF	HMRHM	MHXRTHXT
ZJZBZJFJ	*QFB	*HXRTTHXT	MXTRRR
QZJFBB	*JJFBBJ	*HXHRHHT	MHXRHMR
QZJBQZBZ	QJB	*HMRHMTH	*MXRTTHMM
*ZQBFQJFZ	ZQBZZZZZ	*MHXRXMHR	HMRTTHMR
ZQBZQFB	ZQBZJF	MXRMXTR	MXRHMT
QJBZJFBB	*QJBZFZQB	*HXTRH	*XXTRRX
*QJBZQBZB	ZJZBQ	*HMRHMTXR	*HMRXHHH
*ZJBFFZJF	QJBFFZJF	MHXRMXTR	MHXRHHM
*ZZJBFQJB	*ZBZBZQBZ	*MXRHMRHR	HMTRRX
*JQBQJFBB	ZQBFFZQB	HMRTHMR	HMR
*ZJZFQJB	ZJZBFZQB	*MXRTHHT	MXRH
*QBJBQZBQ	QZJF	*TMR	*RHXRMHXR
*ZJFBZ	*QZJFQBBJ	HMRTMXT	HXHRHH
QJBFQJFB	ZJZFBBJ	*MHXRRMTR	HXHRTHXT
QJBZQBZQ	ZQBQZBZQ	*TXRMXR	*MXRTHHHM
QJFBBB	*QJBZJQ	MHXRHXT	MHXRTMHR
*FJBQJB	*QJBFZZF	*MXRTTMXX	*MXTT
QZJBFQZB	*QFJFBB	*XMRMXTRR	*HXHRHXRH
*ZQBZQFJB	ZQBFQJF	MHXRH	HXHRHXTX
QZJBZ	*ZQBZQFZ	*MHRTX	HMRHMTX
*BZJBQZJB	ZQFBBJ	*HMRTMXTH	HMRHMTR
*QQBZ	*QZBFJ	*HXMRM	*HMTRMRRX

(Appendix C continues)

Appendix C (continued)

Stimuli presented to subjects for Experiment 2

Testing Phase Stimuli (* indicates Nongrammatical item)

Transfer Items		Control Items	
QJBZQF	*QJBJQBQ	*HXHHRRR	HXHT
*ZQBJZZZ	ZQBFZQB	*MXRHXM	HMRMHRHM
ZJZF	QZJBZZQ	*HXHRHMXT	MHXT
QZJBZJZB	*ZQFBQBBJ	*MTRHHM	*MHXMR
*ZJZBZQJF	*QZBBQJB	MXRTHXT	HXHRTHMR
ZJFBBBJ	QJBFFQJB	HXTRRRX	*MHXTMRRX
*ZQBQBF	ZJZBZZ	*MTR	MHXTRR
*ZJZZBBB	*QJBFFZQQ	HMRTTMXT	HMRHXT
*ZJQBQ	*QZJBJQZB	*MXRHTHMR	*MMRHHHM
ZQBZQFJ	*QZJBBQFB	MXRTTHXT	*HHHRMXRH
ZJZBQZJB	*QQBZZZQ	*HXHRHHXM	*HMRTRXT
QZJBQJFB	*ZJZBZZF	MHXRHXHR	*HRHRHMRH
QJBQJFBB	*ZJZBZZJQ	*HRHT	*MMRH

Appendix D

Stimuli presented to subjects for Experiment 3

Learning Phase	Testing Phase (* indicates a Nongrammatical item)	
TSSXXHPS	TXXTTHPS	*HSTXHHS
TSXXTHH	PTTTTTTHH	*PTTTHPHS
PHPXHPS	TSXXTHPS	*PTHHHH
TSSSXXHH	*PTHPXHSP	TSXXHH
TSXS	*HPXTHH	TPHH
PTTHH	PHPXHH	TSXXTTHH
TXXTHPS	PHPXTTHH	*TTHH
PHPXTHPS	TXXHH	TXXTTHH
PTHPS	TXXTHH	*TXXTHPT
PTTHPS	*TXXHX	TSSSSXS
PHH	TSSXXTHH	*TSSXXHSS
TXXHPXXHH	*PHTTTHH	*PXPHXHTT
PHPXHH	*SHPXTHH	*TSXXPH
TSXXTHPS	PHH	PTTTHH
TXS	*PTHPPPS	TSXXHPS
TSSXXHH	PTHPXHPS	*TXHPS
TXXTTTHH	*TXH	*PTTPS
TXXXTHPS	*PHXPHXPX	PTTTTHPS
PTHPXHH	*PHTHH	PTTHPS
TSSSXS	*PTTTHT	TSSXXHPS
	*TXPH	*XXSHT
	PHPXTHPS	*PSXS
	*SXXHPS	TSSXS
	PHPS	PTTHPXHH
	TXS	*TPTXS

Appendix E

Stimuli presented to subjects for Experiment 4

For visual stimuli, the numbers 1-5 indicate locations, left to right, in 1 of 5 boxes located horizontally across a computer screen. Auditory stimuli are represented by 5 tones ranging from very low pitch (1) to very high pitch (5).

Learning Phase	Testing Phase (* indicates a Nongrammatical item)	
144	3144	32252
3252	*13334142	*14344
32553412	*32255422	3222252
352	*3553413	13415412
3553412	*13441112	*415344
141544	32255412	355344
14153412	14153412	*15145433
1333412	141544	*3514
13412	133412	3553344
13344	35533412	*3344
32255412	*13312	*14333
3255344	3255412	*133343
35533344	325544	*1252
1415412	*31352	1412
32225544	14153344	32553412
35533412	35544	32553344
1341544	*354	13333344
35541544	144	*325514
325544	*4235442	32255344
322252	133344	352
	*134444	*35412
	13341544	*35545
	*13415421	13333412
	*2415344	*55243
	*14514515	*255412

References

Alba, J.W., & Hasher, L. (1983). Is memory schematic? Psychological Bulletin, 93, 203-231.

Anderson, J.R. (1983). The architecture of cognition. Cambridge, MA: Harvard University Press.

Anderson, J.R. (1985). Cognitive psychology and its implications (2nd ed.). New York: W.H. Freeman and Company.

Bartlett, F.C. (1932). Remembering: A study in experimental and social psychology. Cambridge, England: Cambridge University Press.

Becker, C.A. (1976). Allocation of attention during visual word recognition. Journal of Experimental Psychology: Human Perception and Performance, 2, 556-566.

Berry, D.C., & Broadbent, D.E. (1988). Interactive tasks and the implicit-explicit distinction. British Journal of Psychology, 79, 251-272.

Brooks, L.R. (1978). Nonanalytic concept formation and memory for instances. In Rosch, E., & Lloyd, B.B. (eds.), Cognition and categorization, pp. 169-211. Hillsdale, NJ: Lawrence Earlbaum Associates.

Brooks, L.R., & Vokey, J.R. (1991). Abstract analogies and abstracted grammars: A comment on Reber (1989), and Mathews, et al. (1989). Journal of Experimental Psychology: General, 120, 316-323.

Bruner, J.S., Goodnow, J., & Austin, G. (1956). A study of thinking. New York: Wiley.

Chang, T.M. (1986). Semantic memory: Facts and models. Psychological Bulletin, 99, 199-220.

Clarke, R., & Morton, J. (1983). Cross-modality facilitation in tachistoscopic word recognition. Quarterly Journal of Experimental Psychology, 35A, 79-96.

Cleeremans, A., & McClelland, J.L. (1991). Learning the structure of event sequences. Journal of Experimental Psychology: General.

Collins, A.M., & Loftus, E.F. (1975). A spreading-activation theory of semantic processing. Psychological Review, 82, 407-428.

Collins, A.M., & Quillian, M.R. (1969). Retrieval time from semantic memory. Journal of Verbal Learning and Verbal Behavior, 8, 240-247.

Dulany, D.E., Carlson, R.A., & Dewey, G.I. (1984). A case of syntactical learning and judgment: How conscious and how abstract? Journal of Experimental Psychology: General, 113, 541-555.

Ellis, H.C. (1965). The transfer of learning. New York: Macmillan.

Fodor, J.A., & Pylyshyn, Z.E. (1988). Connectionism and cognitive architecture: A critical analysis. Cognition, 28, 3-71.

Fong, G.T., & Nisbett, R.E. (1991). Immediate and delayed transfer of training effects in statistical reasoning. Journal of Experimental Psychology: General, 120, 34-45.

Forster, K.I. (1981). Priming and the effects of sentence and lexical contexts on naming time: Evidence for autonomous lexical processing. Quarterly Journal of Experimental Psychology, 33A, 465-495.

Gentner, D. (1983). Structure mapping: A theoretical framework for analogy. Cognitive Science, 7, 155-170.

Gick, M.L., & Holyoak, K.J. (1980). Analogical problem solving. Cognitive Psychology, 12, 306-355.

Gick, M.L., & Holyoak, K.J. (1983). Schema induction and analogical transfer. Cognitive Psychology, 15, 1-38.

Glass, A.L., & Holyoak, K.J. (1974/1975). Alternative conceptions of semantic theory. Cognition, 3, 313-339.

Graf, P., & Ryan, L. (1990). Transfer-appropriate processing for implicit and explicit memory. Journal of Experimental Psychology: Learning, Memory, and Cognition, 6, 978-992.

Graf, P., Shimamura, A.P., & Squire, L.R. (1985). Priming across modalities and priming across category levels: Extending the domain of preserved function in amnesia. Journal of Experimental Psychology: Learning, Memory, and Cognition, 11, 386-396.

Hanson, S.J., & Burr, D.J. (1990). What connectionist models learn: Learning and representation in connectionist networks. The Behavioral and Brain Sciences, 13, 471-518.

Hayes, N.A., & Broadbent, D.E. (1988). Two modes of learning for interactive tasks. Cognition, 28, 249-276.

Hintzman, D.L. (1986). "Schema abstraction" in a multiple-trace memory model. Psychological Review, 93, 411-428.

Holland, J.H., Holyoak, K.J., Nisbett, R.E., & Thagard, P.R. (1986). Induction: Processes of inference, learning, and discovery. Cambridge, MA: MIT Press.

Howard, Jr., J.H., & Ballas, J.A. (1982). Acquisition of acoustic pattern categories by exemplar observation. Organizational Behavior and Human Performance, 30, 157-173.

Jacoby, L.L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. Journal of Experimental Psychology: General, 110, 306-340.

Katona, G. (1940). Organizing and memorizing. New York: Columbia University Press.

Kirsner, K., Milech, D., & Standen, P. (1983). Common and modality-specific processes in the mental lexicon. Memory and Cognition, 11, 621-630.

Laird, J.E., Newell, A., & Rosenbloom, P.S. (1987). SOAR: An architecture for general intelligence. Artificial Intelligence, 33, 1-64.

Lewicki, P., Czyzewska, M., & Hoffman, H. (1987). Unconscious acquisition of complex procedural knowledge. Journal of Experimental Psychology: Learning, Memory, and Cognition, 13, 523-530.

Lewicki, P., Hill, T., & Bizot, E. (1988). Acquisition of procedural knowledge about a pattern of stimuli that cannot be articulated. Cognitive Psychology, 20, 24-37.

Manza, L. (1991). Modality effects and implicit learning: Transfer and abstract representations of complex information. Unpublished master's thesis, Brooklyn College of the City University of New York, Brooklyn, NY.

Manza, L. (1992). Transfer-appropriate processing and the representation of implicitly acquired information. Paper presented at the 63rd Annual Meeting of the Eastern Psychological Association, Boston, MA.

Mathews, R.C. (1990). Abstractness of implicit grammar knowledge: Comments on Perruchet and Pacteau's analysis of synthetic grammar learning. Journal of Experimental Psychology: General, 119, 412-416

Mathews, R.C. (1991). Abstractness of implicit grammar knowledge: Comments on Perruchet and Pacteau's analysis of synthetic grammar learning. Journal of Experimental Psychology: General, 119, 412-416

Mathews, R.C., Buss, R.R., Stanley, W.B., Blanchard-Fields, F., Cho, J.R., & Druhan, B. (1989). Role of implicit and explicit processes in learning from examples: A synergistic effect. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15, 1083-1100.

McAndrews, M.P., & Moscovitch, M. (1985). Rule-based and exemplar-based classification in artificial grammar learning. Memory and Cognition, 13, 469-475.

McClelland, J.L., & Rumelhart, D.E. (1985). Distributed memory and the representation of general and specific information. Journal of Experimental Psychology: General, 114, 159-188.

McCloskey, M., & Glucksberg, S. (1979). Decision processes in verifying category membership statements: Implications for models of semantic memory. Cognitive Psychology, 11, 1-37.

Medin, D.L. (1989). Concepts and conceptual structure. American Psychologist, 44, 1469-1481.

Medin, D.L., & Schaffer, M.M. (1978). Context theory of classification learning. Psychological Review, 85, 207-238.

Minsky, M. (1975). A framework for representing knowledge. In Winston, P.H. (ed.), The psychology of computer vision. New York: McGraw Hill.

Morton, J. (1969). Interaction of information in word recognition. Psychological Review, 76, 165-178.

Nosofsky, R.M., Kruschke, J.K., & McKinley, S.C. (1992). Combining exemplar-based category representations and connectionist learning rules. Journal of Experimental Psychology: Learning, Memory, and Cognition, 18, 211-233.

Novick, L.R. (1990). Representational transfer in problem solving. Psychological Science, 1, 128-132.

Perruchet, P., Gallego, J., & Savy, I. (1990). A critical reappraisal of the evidence for unconscious abstraction of deterministic rules in complex experimental situations. Cognitive Psychology, 22, 493-516.

Perruchet, P., & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge. Journal of Experimental Psychology: General, 119, 264-275.

Perruchet, P., & Pacteau, C. (1991). Implicit acquisition of abstract knowledge about artificial grammar: Some methodological and conceptual issues. Journal of Experimental Psychology: General, 120, 12-116.

Pikas, A. (1966). Abstraction and concept formation. Cambridge, MA: Harvard University Press.

Reber, A.S. (1967). Implicit learning of artificial grammars. Journal of Verbal Learning and Verbal Behavior, 7, 317-327.

Reber, A.S. (1969). Transfer of syntactic structure in synthetic languages. Journal of Experimental Psychology, 81, 115-119.

Reber, A.S. (1989). Implicit learning and tacit knowledge. Journal of Experimental Psychology: General, 118, 219-235.

Reber, A.S. (1990). On the primacy of the implicit: Comment on Perruchet and Pacteau. Journal of Experimental Psychology: General, 119, 340-342.

Reber, A.S., & Allen, R. (1978). Analogic and abstraction strategies in synthetic grammar learning: A functionalist interpretation. Cognition, 6, 189-221.

Reber, A.S., & Lewis, S. (1977). Implicit learning: An analysis of the form and structure of a body of tacit knowledge. Cognition, 5, 333-361.

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.

Reder, L.M., Anderson, J.R., & Bjork, R.A. (1974). A semantic interpretation of encoding specificity. Journal of Experimental Psychology, 102, 648-656.

Reed, S.K. (1987). A structure-mapping model for word problems. Journal of Experimental Psychology: Learning, Memory, and Cognition, 13, 124-139.

Roediger, H.L., & Blaxton, T.A. (1987). Effects of varying modality, surface features, and retention interval on priming in word-fragment completion. Memory & Cognition, 15, 379-388.

Rosch, E. (1978). Principles of categorization. In Rosch, E., & Lloyd, B.B. (eds.), Cognition and categorization, pp. 28-48. Hillsdale, NJ: Lawrence Earlbaum Associates.

Roussel, L.G., Mathews, R.C., & Druhan, B.B. (1990). Rule induction and interference in the absence of feedback: A classifier system model. Proceedings of the Twelfth Annual Conference of the Cognitive Science Society, Cambridge, MA, 733-740.

Rumelhart, D.E. (1989). The architecture of the mind: A connectionist approach. In Posner, M.I. (ed.). Foundations of Cognitive Science, pp. 133-159. Cambridge, MA: MIT Press.

Schacter, D.L. (1992). Forms of memory: Cognitive and neuropsychological perspectives. Invited address at the 63rd Annual Meeting of the Eastern Psychological Association, Boston, MA.

Schacter, D.L., & Graf, P. (1989). Modality specificity of implicit memory for new associations. Journal of Experimental Psychology: Learning, Memory, & Cognition, 15, 3-12.

Schank, R.C., & Abelson, R. (1977). Scripts, plans, goals, and understanding. Hillsdale, NJ: Lawrence Earlbaum Associates.

Servan-Schreiber, E., & Anderson, J.R. (1990). Learning artificial grammars with competitive chunking. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16, 592-608.

Singley, M.K., & Anderson, J.R. (1989). The transfer of cognitive skill. Cambridge, MA: Harvard University Press.

Smith, E.E. (1989). Concepts and induction. In Posner, M.I. (ed.). Foundations of Cognitive Science, pp. 501-526. Cambridge, MA: MIT Press.

Smith, E.E., & Medin, D.L. (1981). Categories and concepts. Cambridge, MA: Harvard University Press.

Smith, E.E., Shoben, E.J., & Rips, L.J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. Psychological Review, 81, 214-241.

Stanovich, K.E., & West, R.F. (1981). The effect of sentence context on ongoing word recognition: Tests of a two-process theory. Journal of Experimental Psychology: Human Perception and Performance, 7, 658-672.

Thorndike, E.L., & Woodworth, R.S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. Psychological Review, 8, 247-261.

Thorndike, E.L. (1922). The effect of changed data upon reasoning. Journal of Experimental Psychology, 5, 33-38.

Thorndike, E.L. (1925). The psychology of learning. New York: Columbia University Press.

Vokey, J.R., & Brooks, L.R. (1992). Salience of item knowledge in learning artificial grammars. Journal of Experimental Psychology: Learning, Memory, and Cognition, 18, 328-344.

Wertheimer, M. (1945). Productive thinking. New York: Harper.